

The Unequal Availability of Rental Housing Information Across Neighborhoods

Max Besbris, Ariela Schachter, and John Kuk

ABSTRACT As more urban residents find their housing through online search tools, recent research has theorized the potential for online information to transform and equalize the housing search process. Yet, very little is known about what rental housing information is available online. Using a corpus of millions of geocoded Craigslist advertisements for rental housing across the 50 largest metropolitan statistical areas in the United States merged with census tract-level data from the American Community Survey, we identify and describe the types of information commonly included in listings across different types of neighborhoods. We find that in the online housing market, renters are exposed to fundamentally different types of information depending on the ethnorracial and socioeconomic makeup of the neighborhoods where they are searching.

KEYWORDS Rental housing • Housing search • Craigslist • Neighborhood inequality • Racial/ethnic inequality

Introduction

Residential mobility decisions, which are predicated on information about available housing units and accessible neighborhoods, critically shape life chances (Bischoff and Owens 2019; Chetty and Hendren 2018; Sampson 2012; Sharkey and Faber 2014) and rates of residential segregation (Krysan 2002; Krysan and Crowder 2017; Massey and Denton 1993; South et al. 2011). Long-standing interest in variations in the availability of housing information across ethnorracial groups (e.g., Courant 1978) has culminated in recent insights regarding how homeseekers form their choice sets and make decisions about where to move (Bader and Krysan 2015; Bruch and Swait 2019; Havekes et al. 2016; Krysan and Crowder 2017). Broadly, this research is committed to the idea that differential housing outcomes by race/ethnicity—such as those previously found in research on residential mobility (e.g., Bruch and Mare 2006; Bruch and Swait 2019; Crowder and South 2005; Logan and Alba 1993)—are both a product and cause of sociospatial inequality (Krysan and Crowder 2017).

In this article, we gather data and employ methods rarely used by demographers to further research on residential mobility in the United States in two key, related ways. First, in light of the rapidly changing housing search process, we examine what kind of

information about rental housing is available online. The most recent American Housing Survey (2017) found that housing websites are now a primary source of information for all urban homeseekers. Despite the turn toward understanding the sources of information homeseekers rely on when making mobility decisions and the expanding use of online search tools, we lack a thorough description of the housing market information that is readily available online. We examine whether this information is similar across neighborhoods or whether, like other sources of information, housing websites present segmented and segregated information that tracks with racial/ethnic and socioeconomic forms of sociospatial inequality. Doing so helps adjudicate between perspectives that express some optimism about the potential for online search tools to reduce racial information inequalities in residential mobility decision making (e.g., Krysan and Crowder 2017; McLaughlin and Young 2018; Palm and Danis 2001) and those that argue that any new technology that fosters mobility will likely reproduce existing inequalities (e.g., Brannon 2017; Massey 2005; Stiel and Jordan 2018). Second, in analyzing an increasingly important source of information for homeseekers, we emphasize that mobility decisions are partly a product of the supply of information on available units.

To understand the supply of information, we collected 1.6 million geocoded advertisements for rental housing from the 50 largest metropolitan areas in the United States posted on Craigslist, the dominant platform for today's metropolitan rental housing market (Boeing 2020; Boeing and Waddell 2017).¹ Using computational text analysis techniques, we first identify common types of information displayed in online housing advertisements. Next, we demonstrate that advertisements for rental housing largely reflect existing sociospatial inequalities: the information about available housing units varies depending on the surrounding neighborhood's ethnoracial makeup and rate of households in poverty. Listings in poorer neighborhoods tend to focus on tenant qualifications, such as financial requirements and lack of eviction or criminal history, rather than describing the housing unit's amenities. But even among non-poor neighborhoods, listings in Black and Latino neighborhoods focus disproportionately on tenant (dis)qualifications compared with listings in otherwise similar White neighborhoods, underscoring the racialized nature of information in the online rental housing market. In contrast, listings in White and Asian neighborhoods—regardless of poverty level—are more likely to describe the aesthetic qualities of housing units. Finally, listings with higher asking rents in White and Asian neighborhoods, particularly those with higher poverty rates and thus more gentrification potential (Hwang 2015, 2016; Hwang and Sampson 2014; Timberlake and Johns-Wolfe 2017), are more likely to describe desirable neighborhood characteristics. These differences highlight the importance of studying the information environment itself (Bruch and Feinberg 2017). In the online housing market, renters are exposed to fundamentally different types of information depending on the ethnoracial and socioeconomic makeup of the neighborhoods in which they search. These information differences may attract or repel certain types of homeseekers and reify place reputations—key mechanisms of residential sorting that operate as both outcomes and causes of sociospatial inequality (Krysan and Crowder 2017).

¹ Craigslist operates as a classifieds website outside of the United States as well, but it is not the primary housing website in many other countries (see Rae 2015).

Information and Residential Mobility in the Rental Market

Since the Great Recession, a growing number of American households have become renters, and the rental market continues to be where the majority of African Americans, Latinos, and immigrants find their housing (Ellen and Karfunkel 2016; National Multifamily Housing Coalition 2016; Schachter and Besbris 2017). Renters, compared with homeowners, have higher rates of residential mobility, different rates of racial/ethnic segregation, and distinct choice constraints in their residential mobility decisions; further, most metropolitan renters face a market with higher demand relative to supply as well as rising costs (DeLuca et al. 2013; Desmond and Shollenberger 2015; Friedman et al. 2013; Joint Center for Housing Studies 2019; Pilkauskas and Michelmore 2019). Landlords are therefore well positioned to exploit information asymmetries, potentially shaping how renters are sorted (Garboden and Rosen 2019; Rosen 2014).

As a growing number and proportion of Americans have become renters and rental housing affordability has declined, the housing market in general—and the market for rental housing in particular—increasingly operates online. Recent survey data show that, across racial/ethnic groups, internet sites like Craigslist are one of the two most common ways (along with word of mouth) that homeseekers in urban areas find places to live (American Housing Survey 2017).² In short, housing websites—in conjunction with the sharp decline in unequal access to the internet in urban areas (Anderson and Perrin 2021)—are transforming residential search and mobility processes (see Schachner and Sampson 2020:679).

Online rental housing advertisements are a point of connection for landlords, who control the supply of rental housing, with prospective renters, whose preferences shape demand. By serving as this point of connection, rental housing advertisements can influence the types of households who apply and those who do not. In other words, advertisements for rental housing are a key source of information for renters making residential mobility decisions.

Online listings are also often the first signal that prospective renters receive about whether a particular unit matches their housing preferences, and the iterative and imbricated nature of the housing search process means that prospective tenants learn about potential places to live as they browse listings (Krysan and Crowder 2017). Moreover, online search tools allow for easier and faster comparison across units. On the one hand, this could help equalize searches because a wide range of listings can be accessed. On the other hand, it could also heighten particular signals as searchers quickly screen out a high volume of listings (Bruch et al. 2016).

When individuals begin searching for housing, one of their initial tasks is to determine in which neighborhoods they would consider living (Bader and Krysan 2015). Of course, renters face structural constraints (e.g., price, geography), have preexisting information about neighborhoods, and may have preferences for certain neighbor-

² The use of the internet as a rental housing search tool is differentiated across levels of education. Renters with a bachelor's or advanced degree report finding their housing via the internet at double the rate of renters with a high school diploma or less. However, this trend is reversed for renters with families: these renters are nearly twice as likely to find housing using the internet if they have a high school education or less than if they have a bachelor's or advanced degree.

hoods based on factors such as their social networks and commutes. But this pre-existing information tends to be minimal (Krysan and Bader 2009; Lareau 2014).³ Previous research has shown that certain types of information can affect homeseekers' understandings of different neighborhoods as more or less appropriate places to live. Descriptions of local amenities, for example, influence neighborhood selection, and homeseekers are sensitive to signals about crime and safety—although signals about crime may operate as proxies for ethnoracial makeup (Krysan and Crowder 2017; see also Quillian and Pager 2001). In fact, language that is not overtly racial can still provide cues about a given neighborhood's demographics (Besbris 2016, 2020; Besbris and Faber 2017; Howell and Emerson 2018; Kennedy et al. 2021; Korver-Glenn 2018, 2021) and can subsequently affect residential mobility decisions and economic decisions more generally (Besbris et al. 2015, 2019; Krysan and Crowder 2017). Landlords may similarly be influenced by shared perceptions of neighborhoods when they compose their advertisements, and the information they provide about available housing likely both reflects existing patterns of residential sorting and perpetuates them.

After homeseekers select a neighborhood or set of neighborhoods in which to search, they must then compare available housing units (Krysan and Crowder 2017:53). Variation in descriptions about the units is a key component of the residential selection process (see Harvey et al. 2020; Rosenblatt and DeLuca 2012; Wood 2014). The information on sites like Craigslist is potentially influential for the selection of both neighborhoods and individual housing units: it is far more robust than the types of information homeseekers tend to gather from their social networks (Carrillo et al. 2016; Lareau 2014), and it is updated in real time as renters go through the search process.

Most online housing platforms like Craigslist require landlords to provide the location and price of a listed unit. However, landlords are free to choose what other types of information to include in their advertisements, such as descriptions of the unit. Whether a landlord is motivated by profit or bias, or is simply trying to provide relevant information to prospective tenants, the selective inclusion and exclusion of information in rental housing advertisements may influence homeseekers' residential mobility decisions. In other words, examining the content of online housing advertisements is essential because it reflects perceptions about the types of people who belong in particular neighborhoods, facilitates landlords' selection of particular kinds of renters, and enables renters to select neighborhoods and units that match their preferences.

Data and Methods

Following growing recognition of the value of data collected online for understanding demographic processes (Cesare et al. 2018), we examine advertisements collected from Craigslist. Not all rental housing in the United States is advertised on Craigslist; indeed, Boeing (2020) found that advertisements for rental housing on Craigslist in

³ Additionally, according to the Current Population Survey, 3% to 5% of the U.S. population since 2005 has engaged in cross-county residential mobility annually (Frey 2019). Thus, in a given year, tens of millions of Americans are gathering information on neighborhoods in cities where they do not currently live.

2014 were overrepresented in neighborhoods with higher shares of White residents, demonstrating how offline inequalities are reproduced in the supply of information online. However, our goal is to understand what kind of information is shared on Craigslist, the most comprehensive and timely source of housing market information in the United States (Boeing and Waddell 2017; Kuk et al. 2021).

We designed a set of Python scripts to crawl Craigslist and gather information from rental ads, including listing date, rent (price), square footage and other unit characteristics, neighborhood name, geolocation, and the full text of the advertisement. We include all Craigslist sites that correspond to the 50 largest metropolitan statistical areas (MSAs) in the United States.⁴ Posters creating ads for rental housing are asked to supply the closest cross streets for the listing, and this position is plotted on a Google maps image embedded within the advertisement. We identify each listing's location using the approximate longitude and latitude from the cross-street plot on Google maps. Across our metro areas, 12% of all listings are missing a geocode and are thus excluded from all analyses presented here. We use the geocodes to assign each advertisement to a census tract using the ArcGIS geographic join tool, which returns 15-character Federal Information Processing Standards census tract codes to indicate to which census tract each geocode belongs.⁵ The Python scripts revisit each MSA Craigslist site once per week to check whether each currently posted listing is new, in which case all information is scraped; if the listing is repeated from the previous week, this information is noted in the database.⁶

From late May 2017 through the middle of February 2018, we collected 3,950,558 listings across all 50 MSAs. We eliminate listings missing price information (about 1%), listings with prices higher than \$10,000 per month (about 1%), and listings that are duplicates (about 50%), yielding 1,697,117 unique geocoded listings.⁷ We then merge our data with 2016 American Community Survey (ACS) five-year pooled data on tract racial/ethnic composition, poverty status, and other neighborhood characteristics relevant to rental market dynamics (U.S. Census Bureau 2017). Because of missing data in our various covariates and listings that do not have enough text for topic modeling, our final sample size is 1,692,639.⁸

A large body of work has shown that two dimensions of neighborhoods together structure the housing market overall and residential mobility in particular: ethnoraacial and socioeconomic composition (Adelman 2005; Charles 2006; Clark 1992, 2009; Clark and Morrison 2012; Crowder and South 2008; Gabriel and Spring 2019; Krysan and Crowder 2017; Lee et al. 1994; Sampson and Sharkey 2008; Swaroop and Krysan 2011). Pervasive cross-neighborhood inequality offline motivates us to test

⁴ For most MSAs, the corresponding Craigslist site closely matches census MSA definitions; moreover, because we use only tract-level census data and follow Craigslist market definitions to determine metro area boundaries, any discrepancies do not impact our results (see the online appendix).

⁵ Here, we measure information differences at the tract level. We run similar models at the ZIP-code level in case there are any inaccuracies in matching ads to tracts, and we find similar results (see the online appendix).

⁶ We exclude posts that do not post a date on which the housing will be available because these tend to be spam posts. By scraping weekly, we miss listings that are posted and removed within one week.

⁷ We present regression results based on a sample that includes listings with prices higher than \$10,000 per month in the online appendix. The results are substantively similar to those presented here.

⁸ See the online appendix for tables on the distribution of ads across MSAs and tracts.

whether the entrenched sociospatial hierarchy in U.S. cities is reflected in the types of information contained in housing listings in different types of neighborhoods.⁹ Similar to analyses by Wang et al. (2018), we use 2016 five-year pooled ACS data in most of our analyses to classify tracts into eight different neighborhood types: White nonpoor, White poor, Black nonpoor, Black poor, Latino nonpoor, Latino poor, Asian nonpoor, and Asian poor. Neighborhood racial composition is based on the plurality racial group, and we use a threshold of 30% of tract households living at or below the federal poverty line as our measure of neighborhood poverty. We tested these cutoffs and found no substantive differences using alternative neighborhood classification schemas (results shown in Tables A11 and A14, online appendix). By using a categorical measure of neighborhood type, we can better identify how neighborhood racial composition and socioeconomic status (SES) intersect. We find substantively similar results using continuous measures (see Table A15, online appendix). In additional analyses, we include posted unit rental price and a broader set of neighborhood measures (from 2016 ACS five-year data) that are commonly correlated with neighborhood race/ethnicity and poverty rate, including the proportion of college-educated residents, the proportion of the population that is foreign-born, the proportion of units that are renter occupied, the proportion of units built after 2010, and the neighborhood vacancy rate.¹⁰ We examine these variables, which measure either the quality of the units in a neighborhood or neighborhood demographics, because they could plausibly affect the ways landlords list their units. For example, a more highly educated renter pool could prompt landlords to advertise certain types of amenities, whereas a higher vacancy rate might create more competition for potential renters and prompt landlords to add more information to their listings (Boeing et al. 2020).

Recent work suggests that information presented in online housing advertisements indeed varies by neighborhood. For example, Craigslist advertisements from poorer neighborhoods and neighborhoods with more Black residents contain fewer words, on average, and are less likely to state an exact address for the rental (Boeing et al. 2020). In addition, the prevalence of Craigslist listings relative to the underlying housing stock is greater in neighborhoods with more White, higher-income, and more educated residents (Boeing 2020). Critically, however, little work has examined the main textual description of listings, with the exception of Kennedy et al.'s (2021) recent study. Although collecting and analyzing these data is more time consuming and labor intensive, it allows for a deeper understanding of the information commonly included in online rental housing advertisements. As we will demonstrate, the text of advertisements tends to contain key details about housing and neighborhoods, which likely influence prospective renters' residential mobility decisions. Moreover, unlike the so-called checkboxes analyzed in prior work, which offer landlords limited options/flexibility in describing their units (Boeing et al. 2020), landlords have full discretion to include (or exclude) any type and amount of text-based information they consider necessary to attract their desired tenants in the main listing. Exploring

⁹ The nonrandom selection of neighborhoods by economic conditions during the residential mobility process and the subsequent effect of these conditions on residents also exists outside of the United States (McAvay 2018; van Ham et al. 2018).

¹⁰ These additional measures are meant to further contextualize the main findings on differences in advertisements by neighborhood type (see Varian 2014).

this discretionary information is critical given research in other settings showing that unequal and discriminatory treatment can be more prevalent in contexts where gatekeepers have more discretion (Pager and Shepherd 2008).

We use two computational text analysis approaches to describe the kinds of information available in online housing advertisements and to test for variation across neighborhoods. First, we use structural topic models (STM) to identify common topics or themes in advertisements. Topics are sets of words that frequently co-occur. For example, we might anticipate that rental housing listings are likely to include language about the number of bedrooms and bathrooms. If basic descriptions of the housing unit are a common type of information, topic model results will include a topic with words often found in these descriptions (e.g., “bedroom,” “bathroom”). Topic models do not require researchers to know beforehand which themes will emerge; rather, they take a purely inductive approach by identifying commonly co-occurring sets of words. The researcher then examines the collections of words and identifies their substantive meaning. Thus, topic models allow us to characterize the different content areas that are commonly included in online housing listings without being influenced by any prior assumptions. In addition to providing a description of information commonly presented in online housing advertisements, STM can also be used to compare the prevalence of topics and specific word choices within a topic across different types of neighborhoods (DiMaggio et al. 2013; Roberts et al. 2014). In other words, we can use STM to estimate which types of topics landlords are more likely to use in advertisements for housing in different types of neighborhoods and to compare the types of words used across neighborhoods within the same topic.

To run our STM analysis, we first create a document-term matrix containing information specifying how many times each term appears in each individual advertisement. In preprocessing the corpus, we first convert capital letters to lowercase; remove numbers, stop words (e.g., “a,” “an,” “the,” “am,” “are,” “is”), and punctuation; and conduct stemming to obtain more informative outcomes. Second, we remove low-frequency words—those appearing in less than 1% of ads—which is standard and a crucial step to reduce noise in the outcome (Mosteller and Wallace 1963).

After preprocessing and creating a document-term matrix, we run an STM with seven topics.¹¹ Our STM analysis occurs over three stages. First, we compute topic proportions by each document.¹² When we conduct STM, we include covariates that account for the document-generating process.¹³ However, including covariates in the STM estimation makes minimal difference in the topic model outcomes, and STM without any covariates demonstrates identical results (as shown in the online appendix). Second, we run regression models to estimate the relationship between neighborhood type and topic proportions by including the same set of covariates. We include MSA fixed effects. Finally, we compare word choices within the same topic

¹¹ We follow Chang et al. (2009) in choosing a number of topics that are easily interpretable and convey cohesive meaning. In the online appendix, we present results that use different number of topics.

¹² STM also produces a word-topic matrix, which represents the proportion of word use by each topic.

¹³ Our preferred STM estimation includes the following covariates: the posted unit price, our eight-category neighborhood type classification, the proportion of residents who are college educated, the proportion of units that are renter occupied, the proportion of units that were built after 2014, the proportion of the population that is foreign-born, the vacancy rate, and MSA fixed effects.

between neighborhood types by running a new STM that includes a dummy variable indicating whether the neighborhood is majority White or majority non-White; this type of analysis can be conducted only across two groups, not with our eight-category neighborhood typology.¹⁴

As we describe in detail later, our STM analysis finds substantial variation in the information provided in advertisements across neighborhoods. To further explore these differences, we examine whether individual words are associated with neighborhood characteristics. Focusing only on topics might obscure specific words or phrases such as “Section 8” (a reference to subsidies in the form of vouchers provided to some poor- and moderate-income households for use in the private rental market) or “Whole Foods” (a reference to an upscale grocery store) that vary systematically across different types of neighborhoods. In addition, to implement STM, we have to make decisions in the modeling and interpretation process that might inadvertently influence our findings. Using a second approach to understand patterns in our text data provides a key robustness test.

To identify individual words that tend to distinguish between neighborhoods, we use multinomial inverse regression (MNIR), a powerful tool for measuring how words are associated with continuous measures of neighborhood characteristics.¹⁵ MNIR incorporates high-dimensional data, such as text data with many covariates, into statistical analysis to uncover the strength of correlation between each word and our covariate of choice. For example, MNIR estimates how each word in the corpus is correlated with the poverty rate of the neighborhood where the listing is posted.¹⁶ If “section” (as in “Section 8 housing”) is strongly correlated with listing tract poverty rate, then the MNIR coefficient for “section” will be high. (Because MNIR estimates the strength of association better with continuous measures, we do not use our neighborhood categorization measure.)¹⁷ MNIR achieves the same goal as ordinary least squares (OLS) while addressing statistical issues in high-dimensional data, such as text data. To run MNIR, we use the same preprocessing techniques as in our STM analysis and prepare the document-term matrix. MNIR will produce a vector of coefficients that contains the strength of correlation between the selected neighborhood covariate and each word.¹⁸

¹⁴ Because word comparisons within topics can be calculated only for binary neighborhood measures, here we classify all neighborhoods as majority White or majority non-White.

¹⁵ We choose MNIR over other machine learning methods such as LASSO or ridge regression because MNIR is specifically developed for application to text data (Gentzkow et al. 2019).

¹⁶ Past work has referenced unpublished research on STM analyses of Craigslist data (see Boeing et al. 2021) or used STM to analyze ads in one city (Kennedy et al. 2020). To our knowledge, however, no research has used MNIR to understand the characteristics of rental market information across neighborhoods.

¹⁷ Given that our research questions are ones of association instead of prediction (see Grimmer 2012), log odds and model-based approaches, such as MNIR, are effective for identifying distinct words (Manning et al. 2008; Taddy 2013). However, we find that mutual information models (available upon request) and MNIR produce very similar results.

¹⁸ MNIR produces this vector of covariates by assuming that the document-term matrix is a collection of draws from a multinomial distribution and inverting the regression framework by putting the high dimensional document-term matrix on the left side.

Finally, we conduct ancillary analyses (presented in the online appendix) to further quantify information differences across neighborhoods. In these analyses, we use OLS to identify variation in numeric characteristics of online advertisements, including the overall number of words included in each advertisement and the number of pictures. These analyses identify clear differences in the type of information included in advertisements depending on the neighborhood in which the rentals are located, further supporting the text-based analyses presented here.

Results

Identifying Information Types

We begin by identifying the types of information commonly included in online housing advertisements. [Table 1](#) details our seven topics, listing their labels, the most common words within these topics, and each topic's prevalence average across the entire set of advertisements.

The *general information* topic focuses on the type of building (e.g., apartment building, duplex, single-family home) and, on average, accounts for about 25% of a listing's content. The *availability* topic includes information on how to contact the landlord and view the unit and accounts for 8.5% of listing content on average. *Unit description*, at 25% of average listing content, includes basic descriptions of the size of the unit (e.g., number of bedrooms and bathrooms). The *pet policy* topic (8.3% of content) covers whether and which types of pets are allowed in the unit. These topics cover basic, perfunctory information that is undoubtedly part of homeseekers' decisions. However, they are less subject to landlords' capriciousness in that they are generally objective characteristics about the unit. Additionally, the landlord can signal much of this information—including the number of bedrooms and bathrooms, the availability of a washer and dryer, and the pet policy—in write-in options and checkboxes that are independent of the advertisement's main text and can be used by homeseekers to filter units when searching on Craigslist. Given that other work more thoroughly explored the use of checkboxes in Craigslist advertisements (Boeing et al. 2020), our focus here is on describing the linguistic content of the text in ads.

We focus our analysis on three topics—*logistics*, *unit amenities*, and *neighborhood amenities*—because of their theoretical importance to the residential decision process (Desmond 2016; Harvey et al. 2020; Krysan and Crowder 2017; Rosen 2017; Rosenblatt and DeLuca 2012; Wood 2014). First, the *logistics* topic captures language about the logistics of applying to rent a unit and desired and undesired renter characteristics, including language about housing vouchers, eviction history, and income and credit score requirements. The *logistics* topic accounts for about 11% of the content in an average listing. Landlords include this type of text in their advertisements to try to influence who will contact them for more information or apply to rent their unit (Rosen 2014). [Figure 1](#) shows two examples of listings from our data that have a high proportion of the logistics topic.

Second, the *unit amenities* topic captures more specific language about optional housing features relative to the *unit descriptions* topic (see [Figure 1](#) for examples).

Table 1 STM topics from Craigslist rental listings

Label	Words ^a	%
General	apart, home, communiti, center, pool, call, offer, locat, fit, bedroom	24.6
Availability	apart, avail, leas, today, price, unit, manag, chang, properti, call	8.5
Unit Description	bedroom, room, floor, kitchen, new, includ, larg, bath, bathroom, month	25.2
Pet Policy	pet, apart, home, communiti, polici, offic, restrict, now, hour, hous	8.3
Logistics	rent, month, applic, will, fee, home, credit, pleas, properti, leas	10.6
Unit Amenities	kitchen, floor, center, applianc, room, communiti, loung, fit, stainless, area	12.7
Neighborhood Amenities	park, apart, locat, downtown, walk, shop, restaur, citi, minut, street	10.2

^a Words displayed have simplified forms because they are all stemmed. Every word from STM will be displayed as a stemmed form.

Moreover, in additional analyses, we confirm that these topics are distinct by identifying different prevalence patterns across neighborhoods (see Table A9 and Figure A1, online appendix). On average, the *unit amenities* topic accounts for about 13% of listing content.

Third, the *neighborhood amenities* topic contains language describing neighborhood characteristics. This topic, which accounts for about 10% of the content in an average listing, includes words about parks, shopping, and restaurants, as well as other information about neighborhood location or resources (see Figure 1).¹⁹

Testing for Variation in Available Information Across Neighborhoods

Next, to test whether each topic is disproportionately prominent in advertisements in certain types of neighborhoods, we treat topic prevalence—that is, the proportion of each advertisement’s words that are dedicated to each topic—as a dependent variable. We use OLS models with MSA-level fixed effects to predict prevalence, and we cluster standard errors by MSA.²⁰ We log transform our dependent variables because the distributions of the variables are skewed to the right.²¹ Table 2 reports coefficients predicting the logistics (Model 1a), unit amenities (Model 2a), and neighborhood amenities (Model 3a) topics without additional control measures. We use these models to estimate pairwise differences across neighborhood types by changing the

¹⁹ Both unit amenities and neighborhood amenities potentially vary based on landlords’ advertising strategies and underlying differences in housing and neighborhood characteristics. Our data cannot adjudicate between the two. However, given prior research showing that racial composition tends to predict residents’ and homeseekers’ *assumptions* about unit and neighborhood amenities (Krysan and Crowder 2017), it is unlikely that underlying quality and fixed characteristics of housing and neighborhoods can completely explain differences in ad text.

²⁰ We obtain similar results when we cluster by census tract (see Table A20, online appendix).

²¹ See Figures A2 and A3 (online appendix) for histograms of the dependent variables. The results are similar when we do not log transform our dependent variables (see Table A10 and Figure A4, online appendix).

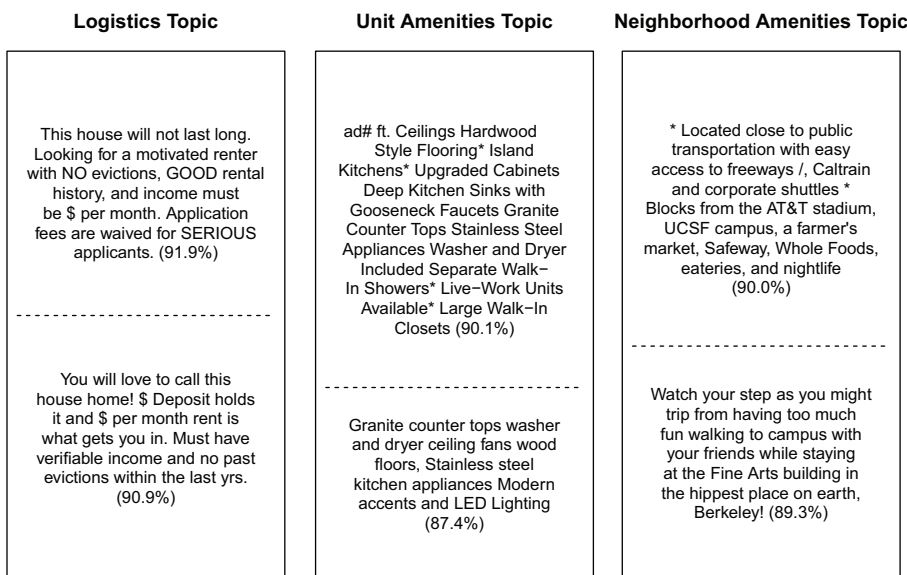


Fig. 1 Example listings of three main topics. The values in parentheses represent the percentage of the listing devoted to the relevant topic. For example, 91.9% of the first listing on the far left of the figure is devoted to the logistics topic.

baseline category in our regression models and generating predicted values. **Figure 2** reports pairwise comparisons of neighborhood types for each of our topics. Because our dependent variables are log-transformed, we present the percentage change in topic proportion, rather than the regression coefficients, for ease of interpretation.

As shown in the first column of **Figure 2**, regardless of race, discussion of rental logistics—income requirements, background checks, renter disqualifications—is almost 50% more prevalent in poorer neighborhoods compared to their same-race, nonpoor counterparts. However, the logistics topic is not just associated with poverty status; with poverty held constant, the logistics topic is 25% to 75% more prevalent in Black and Latino neighborhoods compared with White ones.²² Although discussion of rental logistics is clearly related to neighborhood SES, it is also racialized.

These results suggest that renters searching in higher-poverty, Black, and Latino neighborhoods are more likely to encounter restrictive information in advertisements, which may encourage those searchers who can afford it to select out of the prospective pool of renters before a landlord even begins a formal review of applicants. This language may also lead some renters to believe that these neighborhoods have high rates of crime, eviction, and poverty. That is, this language works in tandem with any preexisting information to further stigmatize neighborhoods (Besbris et al. 2015, 2018, 2019; Sampson and Raudenbush 2004). However, some renters might appreciate having rental requirements presented upfront, and other words within this topic focus on subjects such as lease terms and the rental application process.

²² In these analyses, White, Black, and Latino poor neighborhoods have similar poverty rates (median =38%).

Table 2 Regression results predicting topic proportions with additional neighborhood covariates (dependent variable is log-transformed topic proportion)

	Logistics Topic		Unit Amenities Topic		Neighborhood Amenities Topic	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)
Neighborhood Type						
White poor	0.436** (0.088)	0.520** (0.084)	-0.215* (0.093)	-0.234** (0.078)	0.620** (0.084)	0.382** (0.075)
Black nonpoor	0.487** (0.106)	0.195† (0.104)	-0.594** (0.052)	-0.164** (0.051)	-0.304** (0.054)	-0.036 (0.061)
Black poor	0.995** (0.112)	0.497** (0.122)	-0.792** (0.063)	-0.239** (0.066)	0.193** (0.064)	0.280** (0.064)
Latino nonpoor	0.364** (0.065)	0.362** (0.084)	-0.490** (0.080)	-0.095 (0.072)	-0.124† (0.072)	0.347** (0.055)
Latino poor	0.706** (0.140)	0.664** (0.148)	-0.822** (0.077)	-0.305** (0.079)	0.080 (0.074)	0.478** (0.076)
Asian nonpoor	-0.270* (0.109)	0.236* (0.114)	0.112† (0.058)	-0.081 (0.065)	-0.232** (0.065)	-0.057 (0.092)
Asian poor	0.389 (0.288)	0.802** (0.269)	-0.303 (0.317)	-0.393† (0.216)	0.494** (0.160)	0.439** (0.078)
Unit and Neighborhood Covariates						
Price (\$1,000s)		-0.114** (0.026)		0.336** (0.037)		-0.005 (0.016)
% College		-0.013** (0.002)		0.017** (0.001)		0.018** (0.001)
% Foreign-born		-0.017** (0.004)		0.004* (0.002)		-0.010** (0.002)
% Units renter occupied		-0.009** (0.001)		0.007** (0.001)		0.010** (0.001)
% Unit built after 2010		-0.013* (0.005)		0.047** (0.004)		-0.004 (0.002)
% Vacancy		2.691** (0.269)		-0.911** (0.192)		1.592** (0.212)
MSA Fixed Effects						
Number of Observations	Y 1,692,639	Y 1,692,639	Y 1,692,639	Y 1,692,639	Y 1,692,639	Y 1,692,639
Number of Census Tracts	37,319	37,319	37,319	37,319	37,319	37,319
R ²	.179	.230	.231	.349	.241	.346
Adjusted R ²	.179	.230	.231	.349	.241	.346
Residual SE	1.664	1.611	1.342	1.235	1.104	1.025

Notes: Proportions for the three topics are computed by STM (Roberts et al. 2014). The base category for neighborhood type is White, nonpoor neighborhoods. Listings with rents of more than \$10,000 per month are removed. For neighborhood classification, racial composition is based on the plurality racial group. We use a threshold of 30% of tract poverty rate as our measure of neighborhood poverty. Neighborhood covariates are obtained from 2016 ACS five-year pooled data. Standard errors are clustered at the MSA level.

†*p* < .10; **p* < .05; ***p* < .01

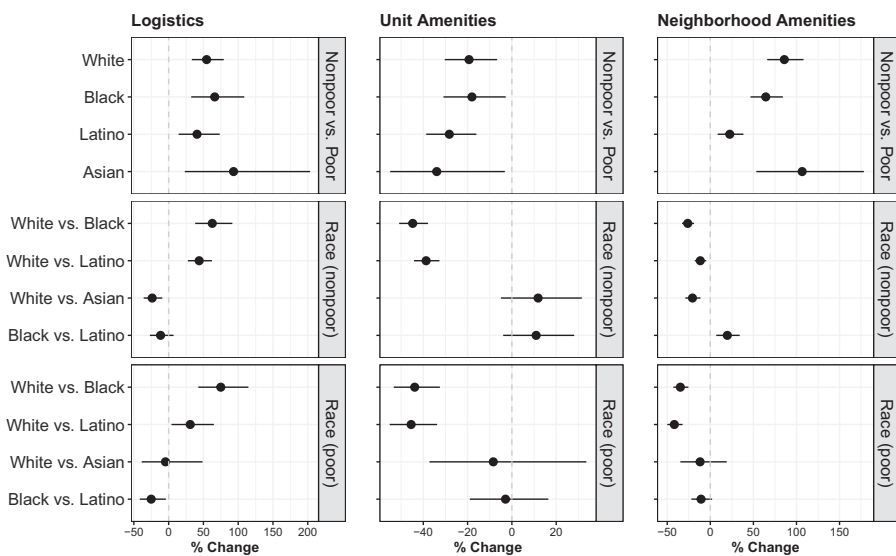


Fig. 2 Pairwise comparison of topic percentages across neighborhood types. The dependent variables are log transformed. The dots and bars indicate, respectively, the percentage change and confidence intervals derived from regression models. For ease of interpretation, we present percentage change instead of regression coefficients. The neighborhood listed first (before “vs.”) is the base category. The first four rows (nonpoor vs. poor) display the regression coefficients for poor neighborhoods when nonpoor neighborhoods for the respective racial group represent the base category. A negative value means that ads in poor neighborhoods provide less information on that topic than those in nonpoor neighborhoods. A positive value means that ads in nonpoor neighborhoods provide less information on that topic than ads in poor neighborhoods. The following eight rows compare differences by racial composition. The first racial group listed is the base category. For example, the coefficient for “White vs. Latino” in the sixth row for the unit amenities topic indicates that Latino nonpoor neighborhoods have ads with 38.7% less content on unit amenities than White nonpoor neighborhoods. Neighborhood racial composition and poverty rate are obtained from 2016 ACS five-year pooled data. The regression models include MSA fixed effects. Standard errors are clustered at the MSA level. The plots are based on the results presented in [Table 2](#).

Although our data do not allow us to test the effects of language on renter behavior, we can examine whether there is any heterogeneity in the choice of words within the logistics topic by neighborhood racial composition. In this analysis, we must use a binary neighborhood classification; we compare word choice in neighborhoods that are majority White with those that are majority non-White. [Figure 3](#) displays the results of this analysis. We find that within the rental logistics topic, the specific words used in advertisements vary along racial lines. Specifically, words such as “must,” “credit,” and “incom[e],” which convey a more exclusionary tone and focus on renter characteristics, predominate in advertisements in majority non-White neighborhoods. In contrast, words in advertisements in majority White neighborhoods convey a more neutral or welcoming tone and a focus on the general rental application process; examples include “will,” “move,” “applic[ation],” “leas[e],” and “free.”

Returning to the second column of [Figure 2](#), we find that advertisements in poorer neighborhoods tend to have less language about unit amenities (e.g., descriptions of building materials, appliances). Relative to same-race, nonpoor neighborhoods, poor

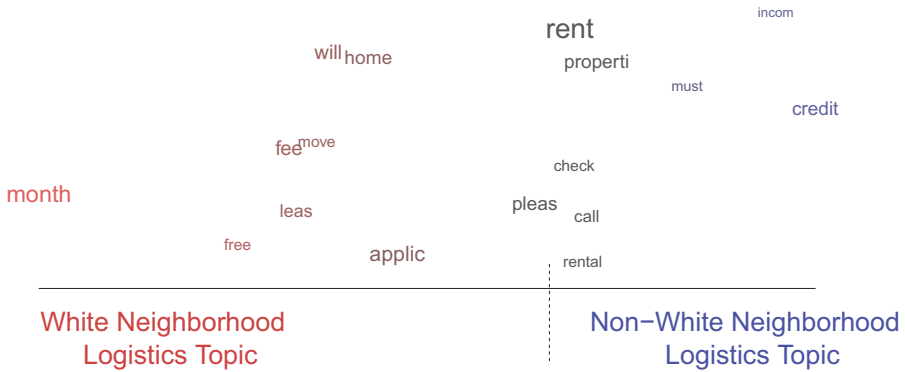


Fig. 3 Choice of words within the logistics topic by neighborhood racial composition. Words that appear on the right side of the figure are more likely to appear in majority non-White neighborhoods. The likelihood of a word's use increases as the word is located farther right. Words that are located to the left of the dashed line are more likely to be used in majority White neighborhoods. The size of the words represents the frequency of the word in the logistics topic.

White, Black, Latino, and Asian neighborhoods all have approximately 20% to 30% less discussion of unit amenities (although the difference for Asian neighborhoods is not significant, perhaps because of small sample sizes). As with the logistics topic, however, we also find a clear racialized pattern of unit amenities language. Among nonpoor neighborhoods, listings in Black and Latino neighborhoods are approximately 40% less likely to contain information about housing unit amenities compared with White neighborhoods. Similarly, listings in poor Black and Latino neighborhoods are more than 40% less likely to contain information about housing unit amenities compared with poor White neighborhoods. In sum, compared with those in White neighborhoods, advertisements in Black and Latino neighborhoods overemphasize renter qualifications and logistics and underemphasize unit amenities.

Discussion of neighborhood amenities (descriptions of nearby parks, restaurants, and so on) displays a more complex pattern. As shown in the final column of [Figure 2](#), listings in poor neighborhoods tend to include *more* discussion of neighborhood amenities relative to listings in neighborhoods of their nonpoor, same-race counterparts. The effect sizes vary from approximately 25% more prevalent in poor Latino neighborhoods to almost 90% more prevalent in poor White neighborhoods (and to 100% more prevalent in poor Asian neighborhoods, although this estimate is very imprecise given our small sample size). Nevertheless, regardless of poverty status, listings in Black and Latino neighborhoods contain less information describing the neighborhood than listings in White neighborhoods with a similar poverty status (both poor and nonpoor). Models 1b, 2b, and 3b in [Table 2](#) repeat this analysis but include additional neighborhood covariates. Although some effect sizes are attenuated, most differences by race and poverty status remain statistically significant.

Why would advertisements for housing in poorer neighborhoods, and particularly in predominantly White and Asian neighborhoods with more college-educated residents, contain more language on neighborhood amenities? This practice may partly reflect landlords' desires to attract higher-SES renters. That is, in trying to attract higher-SES tenants who may be willing and able to pay more rent, landlords are

incentivized to emphasize neighborhood amenities. Many lower-income neighborhoods have some gentrification potential, which could explain why we see more language on neighborhood amenities in poor neighborhoods across racial composition. However, prior research has shown that poor, non-Black neighborhoods are more likely to gentrify than poor, Black neighborhoods (Hwang 2015, 2016; Hwang and Sampson 2014; Timberlake and Johns-Wolfe 2017), which could account for the variance in amenities language across neighborhoods of different racial compositions.

To test whether gentrification potential is driving the prevalence of neighborhood amenities language in poorer neighborhoods, we next examine higher-priced rental units within poorer neighborhoods. We first replicate Table 2 but classify listings based on neighborhood race and whether the listing is above (= high rent) or below (= low rent) the median asking rental price in their metro area (see Table A16, online appendix). We find similar patterns for the logistics and unit amenities topics using the listing rent as our SES measure compared with our typology of neighborhood race by poverty. However, a distinct pattern emerges for neighborhood amenities: although we previously found that neighborhood amenities language is *more prevalent* in poorer neighborhoods, we now find that it is *less* prevalent in lower-rent units. We next interact our binary measure of unit listing rent (high vs. low) with our full, eight-category neighborhood typology, creating 16 neighborhood categories in total (see Table A17 and Figure A6, online appendix).

Even when we account for both neighborhood poverty and rent, large ethnoracial differences remain. With respect to neighborhood amenities, listings in lower-poverty neighborhoods still tend to have less language regarding neighborhood amenities and fewer differences by race or rent. Yet among higher-poverty neighborhoods, higher-rent listings tend to have more neighborhood amenities language; further, regardless of rent, listings in high-poverty White and Asian neighborhoods tend to have more neighborhood amenities language. Altogether, these findings remain consistent with the racialized gentrification processes identified in prior research. Landlords who list higher-rent units in potentially gentrifying neighborhoods tend to include more information about neighborhood amenities in their ads.

Our STM analysis identifies clear differences in information across neighborhoods depending on their racial/ethnic and socioeconomic composition as well as differences in information across units depending on their asking rent. Yet, each topic contains multiple words, and it remains somewhat unclear exactly which words might systematically appear more or less frequently across neighborhoods. Thus, we next use MNIR to identify individual words that are correlated with neighborhood characteristics and to provide a robustness check on our STM findings. We conduct MNIR with continuous neighborhood measures, examining the proportion of residents who are Black, the proportion of residents in poverty, and the proportion of residents who are college-educated. Additional analyses for other neighborhood race and SES measures show similar substantive findings and are reported in the online appendix. Note that, because of MNIR's limitations, these analyses consider just one neighborhood characteristic at a time.

The MNIR results are detailed in Figure 4. For each measure, we list the 50 words that have the highest association with each neighborhood characteristic. Figure 4 displays the words in descending order, from the strongest to the weakest correlation (within this group of relatively highly correlated words). The correlation coefficients



Fig. 4 Words with top 50 correlations with neighborhood covariates. The size of the word is proportional to the coefficient from MNIR. If the size of the word is two times larger than the other one, it means the coefficient is two times higher.

calculated in MNIR have no substantive meaning; MNIR has no standard or accepted cutoffs for what constitutes a weak or strong correlation (unlike Pearson’s *r*). Thus, in MNIR, correlation coefficients for specific words can be interpreted only relative to each other (see the online appendix for MNIR coefficients).

The MNIR results underscore the variation in specific words used by neighborhood racial composition and SES. Beginning with the language associated with a larger proportion of residents who are Black, we see words focused on renter characteristics and (dis)qualifications, such as “evictions,” “section” (short for “Section 8”), “criminal,” and “proof” (of income). This pattern mirrors our findings for the logistics topic in our STM analysis. We also see words about affordability and finances, including “discounts,” “affordable,” “money,” and “income,” although the correlation is not

quite as strong as it was for renter characteristics and (dis)qualifications; for example, “evictions” is about 1.7 times more correlated with the proportion Black than is “income” (see the online appendix). In other analyses (not shown), we confirm that these words are negatively correlated with the proportion of residents who are White.

We see a similar list when we examine words correlated with having a higher neighborhood poverty rate: “evictions,” “criminal” (background), “section” (Section 8), and “proof” (of income) are prevalent, and these words have some of the strongest associations with low-income neighborhoods. We also find similar words about affordability, including “affordable” and “income,” although again “evictions” is about 1.6 times more correlated with the proportion in poverty compared with “income” (see the online appendix). However, we also find that words related to college students—including “campus,” “students,” and “university”—are also highly correlated with neighborhood poverty. College students living in off-campus housing often have little to no personal income, raising neighborhood poverty rates even when they receive familial or other non-income-based financial support (Bishaw 2013). The correlation between neighborhood poverty and these words suggests that landlords in certain higher-poverty neighborhoods might be targeting student renters or that poverty rates in such neighborhoods are high *because* they have many student renters (Ehlenz 2019; Laidley 2014). More broadly, this variation across these first two lists of words underscores the intersectional relationship between neighborhood race and poverty status.

Both of these lists starkly contrast with the words associated with a higher percentage of neighborhood residents with at least a college degree. Rather than mentioning renter qualifications or affordability, advertisements in neighborhoods with more college-educated residents tend to have words describing housing and neighborhood amenities. For example, words such as “rooftop,” “concierge,” “marble,” “elevator,” and “backsplashes” describe housing unit or building amenities that are generally high-end. Other words seem to describe neighborhood or location amenities, such as “whole” and “foods” (as in Whole Foods), “museum,” “nightlife,” and “yoga.” Again, not only do these words appear to be more focused on neighborhood and location characteristics than do the words associated with neighborhoods with a large population percentage Black and in poverty, but they also imply a certain type of neighborhood associated with higher-SES lifestyles and amenities.

The MNIR results offer additional evidence that the content of advertisements depends on a neighborhood’s SES and racial composition. In listings for predominantly Black, Latino, and poor neighborhoods, we find much less evidence of any discussion of the quality/amenities of the housing unit and neighborhood; instead, listings in these neighborhoods focus on affordability and renter qualifications. The emphasis on affordability and renter qualifications likely attracts some prospective renters and repels others, and the lack of emphasis on unit and neighborhood amenities may prevent some prospective renters from considering housing in these neighborhoods. In contrast, in neighborhoods with large proportions of highly educated residents, advertisements do not tend to mention renter qualifications or affordability. The lack of such text does not mean that landlords have no expected or required renter qualifications in these neighborhoods; rather, it seems that landlords who list properties in these neighborhoods do not feel the need to mention them when soliciting renters.

Importantly, certain words are absent in both the MNIR and STM analyses: we find no evidence of explicit racial/ethnic words, perhaps because fair housing laws

and Craigslist posting policies largely prohibit them. Previous studies of overt discrimination in online housing ads have found higher rates of discrimination against renters with children than any other protected category (Oliveri 2010). In addition, various other forms of discrimination, such as racial steering and different response rates to inquiries from different-raced homeseekers, remain prevalent (Besbris 2020; Hanson and Hawley 2011; Hogan and Berry 2011; Korver-Glenn 2021). Yet, the absence of explicit racial/ethnic words in our sample underlines the potential additional influence of the more subtle language differences that we have identified.

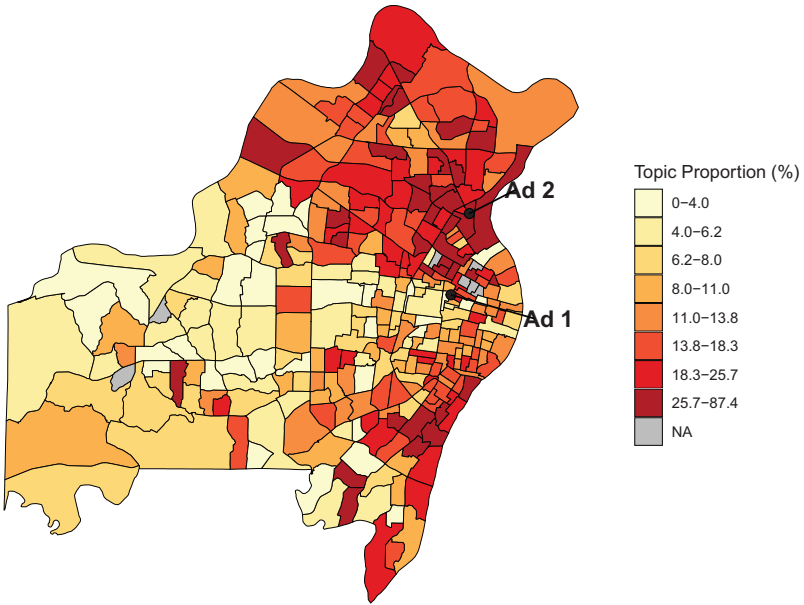
Discussion

Our results highlight the importance of understanding the information environment in which housing searches take place for demographic research on residential mobility. Mobility decisions are predicated on available information, and as shown here, landlords' information-sharing practices do not equalize information across neighborhoods. Multiple analyses reveal that advertisements for units in neighborhoods with more Black, Latino, or poorer residents tend to contain less language describing unit amenities and relatively more language devoted to tenant (dis)qualifications compared with ads from Whiter or lower-poverty neighborhoods. Even in low-poverty Black and Latino neighborhoods, advertisements disproportionately focus on renter (dis)qualifications rather than unit amenities. In contrast, advertisements for housing in White and Asian neighborhoods are more likely to include positive descriptions of neighborhood characteristics; this is particularly true for higher-rent listings in poor White and Asian neighborhoods, which may be undergoing—or poised to undergo—gentrification. Indeed, recent research has demonstrated that the gentrification potential of neighborhoods depends on their existing racial composition (Hwang 2015, 2016). A key limitation of our data is that they are cross sectional. Future research could scrape advertisements longitudinally to pinpoint the relationship between changing information and changing neighborhood demographics, as well as further unpack how other types of variation in neighborhoods (i.e., levels of diversity or segregation) are reflected in advertisements.

To illustrate our findings, Figure 5 presents a pair of maps of neighborhoods less than 6 miles apart in St. Louis, MO. The figure, which shows the prevalence of the logistics topic and rates of Black residents, reveals the sociospatial distribution of information on Craigslist. The maps also contain two dots indicating the location of the rentals described by the advertisements shown in Figure 6.²³ The first ad in Figure 6 is a listing for an apartment in a predominantly White (about 50% of residents), relatively high-poverty neighborhood that has been experiencing upscaling and demographic changes. The listing includes paragraphs describing housing amenities (e.g., “open layout”). It also includes a description of the neighborhood, citing its proximity to universities and to an area with “hustle and bustle.”

²³ These posts were captured after data collection for the results presented here had ended but are qualitatively similar to posts included in our dataset. Both listings were posted within two days of one another during the summer of 2018.

St. Louis Geography of Logistics Topic Prevalence



St. Louis Geography of Black Population

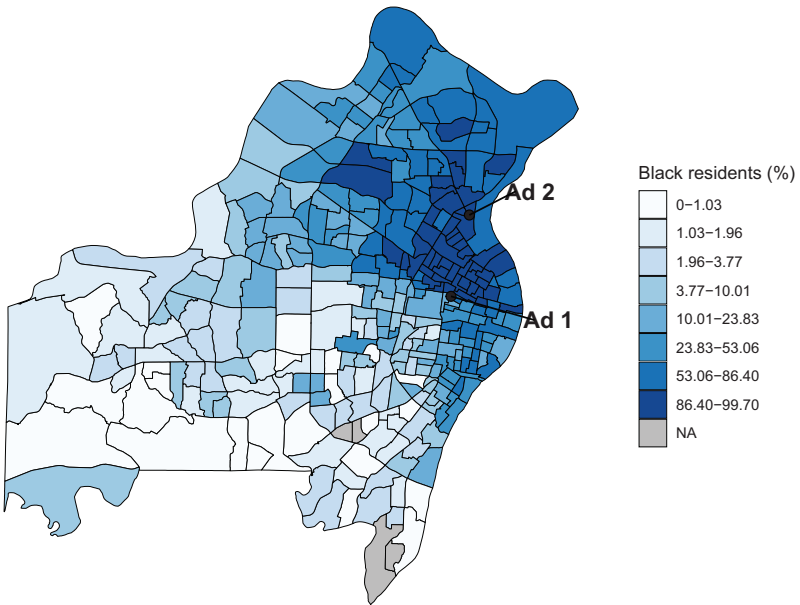
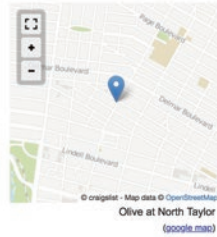


Fig. 5 St. Louis, MO, geography of logistics topic proportion (top) and Black population (bottom)

☆ **\$1300 / 2br - 1150ft² - 2Bedroom Apartment in the CWE with Master Suite (Central West End)**



- 2BR / 2Ba 1150m² available now
- cats are OK - purrr
- dogs are OK - woof
- apartment
- w/d in unit
- detached garage



Welcome to your new home at the historic Lister Building! This wonderful 2-bedroom/2-bath unit in a mid-size apartment building in the Central West End is available second week of June.

The all-electric 1st floor apartment offers a nice open layout with bamboo flooring in the main living areas. The two carpeted bedrooms (one is a master suite) give each other plenty of privacy. The kitchen with granite countertop and plenty of cabinets is every foodie's dream! There is plenty of storage and a full-size washer and dryer are provided.

Garage and/or gated off-street parking available.

Conveniently situated with easy access to major highways, St. Louis University, Washington University and the Cortex, we are located within a short walking distance from the main hustle and bustle of Euclid & Maryland on the corner of North Taylor Avenue & Olive Street.

Contact me to arrange for a showing today!

Is this not quite what you are looking for? No worries, there are other fantastic apartments. Contact me for our upcoming units.

☆ **\$425 / 1br - Affordable 1 Bed Apt! Call Today! (8612 Halls ferry)**



- 1BR / 1Ba available now
- apartment
- laundry on site
- no smoking
- off-street parking
- wheelchair accessible

We have Available Now!

1 Bedroom\$425/mo.

Amenities: On-site laundry, Secure Entry, 24 hour maintenance, some utilities paid by owner.

What you need to bring when you apply:

All applicants must have photos ID, Birth Certificate, Social Security Card and Proof of Income (4 riv Statements) Must be dated in the past 30 to 45 days at time of applying.

- This program is for persons 18 years or older.
- Income must be at least 2 1/2 times the amount of the monthly rent.

- No open Bankruptcies
- No evictions - unless discharged in bankruptcy.
- No outstanding utilities.
- No felonies.

Call Riverview Apartments Today! [show contact info](#).

Fig. 6 Housing advertisement posted from a predominantly White neighborhood with relatively high poverty levels in St. Louis, MO (Ad 1, on the top) and from a predominantly Black neighborhood with relatively high poverty levels in St. Louis, MO (Ad 2, on the bottom)

The second listing in Figure 6 is for an apartment in a predominantly Black neighborhood (about 90% of residents) with relatively high poverty levels. The text in this listing outlines renter qualifications. As prospective renters review listings on Craigslist in the same neighborhood as Ad 1, they will see additional, similar listings containing text referring to housing and neighborhood amenities. In the listings most proximate to Ad 2, prospective renters will see little if any information about the housing or neighborhood amenities but will see a long list of renter disqualifications. Although St. Louis follows the overall trends identified in our sample, future work should test for differences across MSAs based on their levels of segregation and other characteristics to better understand whether and how the racialization of housing information depends on local conditions and histories (Kennedy et al. 2021).

If, as Krysan and Crowder (2017) convincingly argued, most prospective renters are likely to search for housing in neighborhoods they are familiar with through their lived experiences and social networks, then our results suggest that prospective renters receive very different information about units and their surroundings depending on the neighborhoods in which they search. Black and Latino renters, who predominantly search for housing in Black or Latino neighborhoods, encounter strong messaging about their qualifications. Additionally, these differences and the biased spatial availability of advertisements reduce search costs for searchers in Whiter neighborhoods, who are more likely to be White, and expand their mobility options (Boeing 2020).

If at least some prospective (White, nonpoor) renters are open to considering a more heterogeneous group of neighborhoods during the initial stages of their search, the different types of information to which searchers are exposed likely influence their decisions. Language about tenant qualifications, which predominates in Black and Latino neighborhoods, could drive away potential renters who can afford to look elsewhere, particularly because these ads tend to lack text on amenities. More broadly, such information differences contribute to the formation and maintenance of place reputations—acting in concert with homeseekers' existing information to reify perceptions of certain neighborhoods as more or less appropriate for different demographic groups.

Future work should test the effects of language differences described here and explore the myriad potential consequences of our findings for patterns of integration and segregation. We have outlined various ways that these differences matter, but experiments could be used to measure the extent to which homeseekers associate certain types of language with particular neighborhood demographics. Although interest in *how* homeseekers make their residential mobility decisions is growing, more work is needed on the relative importance of various sources of information. In other words, how do homeseekers weigh information found online compared with information from other sources? What is clear from our findings is that online search tools do not serve to erase information differences about available housing across neighborhoods and, as a result, likely foster existing demographic differences in residential mobility. ■

Acknowledgments The authors would like to thank Patrick Ishizuka, Maria Krysan, Barry Lee, Hedwig Lee, Ann Owens, Pat Sharkey, Gerard Torrats-Espinosa, and the anonymous reviewers for comments on previous drafts as well as the Weidenbaum Center at Washington University in St. Louis for financial support.

References

- Adelman, R. M. (2005). The roles of race, class, and residential preferences in the neighborhood racial composition of middle-class Blacks and Whites. *Social Science Quarterly*, 86, 209–228.
- American Housing Survey*. (2017). Retrieved from <https://www.census.gov/programs-surveys/ahs.html>
- Anderson, M., & Perrin, A. (2021). *7% of Americans don't use the internet. Who are they?* [Fact sheet]. Washington, DC: Pew Research Center. Retrieved from <https://www.pewresearch.org/fact-tank/2021/04/02/7-of-americans-dont-use-the-internet-who-are-they/>
- Bader, M. D. M., & Krysan, M. (2015). Community attraction and avoidance in Chicago: What's race got to do with it? *ANNALS of the American Academy of Political and Social Science*, 660, 261–281.
- Besbris, M. (2016). Romancing the home: Emotions and the interactional creation of demand in the housing market. *Socio-Economic Review*, 14, 461–482.
- Besbris, M. (2020). *Upsold: Real estate agents, prices, and neighborhood inequality*. Chicago, IL: University of Chicago Press.
- Besbris, M., & Faber, J. W. (2017). Investigating the relationship between real estate agents, segregation, and house prices: Steering and upselling in New York state. *Sociological Forum*, 3, 850–873.
- Besbris, M., Faber, J. W., Rich, R., & Sharkey, P. (2015). Effect of neighborhood stigma on economic transactions. *Proceedings of the National Academy of Sciences*, 112, 4994–4998.
- Besbris, M., Faber, J. W., Rich, R., & Sharkey, P. (2018). The geography of stigma: Experimental methods to identify the penalty of place. In S. M. Gaddis (Ed.), *Audit studies: Behind the scenes with theory, method, and nuance* (pp. 159–177). New York, NY: Springer.
- Besbris, M., Faber, J. W., & Sharkey, P. (2019). Disentangling the effects of race and place in economic transactions: Findings from an online field experiment. *City & Community*, 18, 529–555.
- Bischoff, K., & Owens, A. (2019). The segregation of opportunity: Social and financial resources in the educational contexts of lower- and higher-income children, 1990–2014. *Demography*, 56, 1635–1664.
- Bishaw, A. (2013). *Examining the effect of off-campus college students on poverty rates* (SEHSD Working Paper No. 2013-17). Washington, DC: U.S. Census Bureau. Retrieved from https://www.census.gov/library/working-papers/2013/acs/2013_Bishaw_01.html
- Boeing, G. (2020). Online rental housing market representation and the digital reproduction of urban inequality. *Environment and Planning A: Economy and Space*, 47, 590–608.
- Boeing, G., Besbris, M., Schachter, A., & Kuk., J. (2021). Housing search in the age of big data: Smarter cities or same old blind spots? *Housing Policy Debate*, 31, 112–126.
- Boeing, G., & Waddell, P. (2017). New insights into rental housing markets across the United States: Web scraping and analyzing Craigslist rental listings. *Journal of Planning Education and Research*, 37, 457–476.
- Brannon, M. M. (2017). Datafied and divided: Techno-dimensions of inequality in American cities. *City & Community*, 16, 20–24.
- Bruch, E., & Feinberg, F. (2017). Decision-making processes in social contexts. *Annual Review of Sociology*, 43, 207–227.
- Bruch, E., Feinberg, F., & Lee, K. Y. (2016). Extracting multistage screening rules from online dating activity data. *Proceedings of the National Academy of Sciences*, 113, 10530–10535.
- Bruch, E. E., & Mare, R. D. (2006). Neighborhood choice and neighborhood change. *American Journal of Sociology*, 112, 667–709.
- Bruch, E., & Swait, J. (2019). Choice set formation in residential mobility and its implications for segregation dynamics. *Demography*, 56, 1665–1692.
- Carrillo, L., Pattillo, M., Hardy, E., & Acevedo-Garcia, D. (2016). Housing decisions among low-income Hispanic households in Chicago. *Cityscape*, 18(2), 109–150.
- Cesare, N., Lee, H., McCormick, T., Spiro, E., & Zaghenni, E. (2018). Promises and pitfalls of using digital traces for demographic research. *Demography*, 55, 1979–1999.
- Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J. L., & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. In Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams, & A. Culotta (Eds.), *Advances in neural information processing systems* (Vol. 22, pp. 288–296). Red Hook, NY: Curran Associates.
- Charles, C. Z. (2006). *Won't you be my neighbor? Race, class, and residence in Los Angeles*. New York, NY: Russell Sage.

- Chetty, R., & Hendren, N. (2018). The impacts of neighborhood on intergenerational mobility I: Childhood exposure effects. *Quarterly Journal of Economics*, *133*, 1107–1162.
- Clark, W. A. V. (1992). Residential preferences and residential choices in a multi-ethnic context. *Demography*, *30*, 451–466.
- Clark, W. A. V. (2009). Changing residential preferences across income, education, and age: Findings from a multi-city study of urban inequality. *Urban Affairs Review*, *44*, 334–355.
- Clark, W. A. V., & Morrison, P. (2012). Socio-spatial mobility and residential sorting: Evidence from a large scale survey. *Urban Studies*, *39*, 3253–3270.
- Courant, P. N. (1978). Racial prejudice in a search model of the urban housing market. *Journal of Urban Economics*, *5*, 329–345.
- Crowder, K., & South, S. J. (2005). Race, class, and changing patterns of migration between poor and non-poor neighborhoods. *American Journal of Sociology*, *110*, 1715–1763.
- Crowder, K., & South, S. J. (2008). Spatial dynamics of White flight: The effect of local and extralocal racial conditions on neighborhood out-migration. *American Sociological Review*, *73*, 792–812.
- DeLuca, S., Garboden, P. M., & Rosenblatt, P. (2013). Segregating shelter: How housing policies shape the residential locations of low-income minority families. *ANNALS of the American Academy of Political and Social Science*, *647*, 268–299.
- Desmond, M. (2016). *Evicted*. New York, NY: Crown.
- Desmond, M., & Shollenberger, T. (2015). Forced displacement from rental housing: Prevalence and neighborhood consequences. *Demography*, *52*, 1751–1772.
- DiMaggio, P., Nag, M., & Blei, D. M. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. *Poetics*, *41*, 570–606.
- Ehlenz, M. M. (2019). The university and its neighborhood: A study of place-making and change. *Journal of Urban Affairs*, *41*, 776–794.
- Ellen, I. G., & Karfunkel, B. (2016). *Renting in America's largest metropolitan areas* (Report of the NYU Furman Center/Capital One National Affordable Housing Landscape). New York, NY: New York University Robert F. Wagner Graduate School of Public Service. Retrieved from https://furmancenter.org/files/NYU_Furman_Center_Capital_One_National_Affordable_Rental_Housing_Landscape_2016_9JUNE2016.pdf
- Frey, W. H. (2019). *For the first time on record, fewer than 10% of Americans moved in a year*. Washington, DC: Brookings Institution. Retrieved from <https://www.brookings.edu/blog/the-avenue/2019/11/22/for-the-first-time-on-record-fewer-than-10-of-americans-moved-in-a-year/>
- Friedman, S., Tsao, H., & Chen, C. (2013). Housing tenure and residential segregation in metropolitan America. *Demography*, *50*, 1477–1498.
- Gabriel, R., & Spring, A. (2019). Neighborhood diversity, neighborhood affluence: An analysis of the neighborhood destination choices of mixed-race couples with children. *Demography*, *56*, 1051–1073.
- Garboden, P. M. E., & Rosen, E. (2019). The threat of eviction: How landlords shape a contingent tenure. *City & Community*, *18*, 638–661.
- Gentzkow, M., Kelly, B. T., & Taddy, M. (2019). Text as data. *Journal of Economic Literature*, *57*, 535–574.
- Grimmer, J. (2012). *Evaluating model performance in fictitious prediction problems*. Discussion of “Multinomial inverse regression for text analysis” by Matthew Taddy. Unpublished manuscript.
- Hanson, A., & Hawley, Z. (2011). Do landlords discriminate in the rental market? *Journal of Urban Economics*, *70*, 99–114.
- Harvey, H., Fong, K., Edin, K., & DeLuca, S. (2020). Forever homes and temporary stops: Housing search logics and residential selection. *Social Forces*, *98*, 1498–1523.
- Havekes, E., Bader, M., & Krysan, M. (2016). Realizing racial and ethnic neighborhood preferences? Exploring the mismatches between what people want, where they search, and where they live. *Population Research and Policy Review*, *35*, 101–126.
- Hogan, B., & Berry, B. (2011). Racial and ethnic biases in rental housing. *City & Community*, *10*, 351–372.
- Howell, J., & Emerson, M. O. (2018). Preserving racial hierarchy amidst changing racial demographics: How neighbourhood racial preferences are changing while maintaining segregation. *Ethnic and Racial Studies*, *41*, 2770–2789.
- Hwang, J. (2015). Gentrification in changing cities: Immigration, new diversity, and racial inequality in neighborhood renewal. *ANNALS of the American Academy of Political and Social Science*, *660*, 319–340.
- Hwang, J. (2016). Pioneers of gentrification: Transformation in global neighborhoods in urban America in the late twentieth century. *Demography*, *53*, 189–203.

- Hwang, J., & Sampson, R. J. (2014). Divergent pathways of gentrification: Racial inequality and the social order of renewal in Chicago neighborhoods. *American Sociological Review*, *79*, 725–751.
- Joint Center for Housing Studies. (2019). *The state of the nation's housing 2019*. Cambridge, MA: President and Fellows of Harvard College. Retrieved from <https://www.jchs.harvard.edu/state-nations-housing-2019>
- Kennedy, I., Hess, C., Paullada, A., & Chasins, S. (2021). Racialized discourse in Seattle rental ad text. *Social Forces*, *99*, 1432–1456.
- Korver-Glenn, E. (2018). Compounding inequalities: How racial stereotypes and discrimination accumulate across the stages of housing exchange. *American Sociological Review*, *83*, 627–656.
- Korver-Glenn, E. (2021). *Race brokers: Housing markets and segregation in 21st century urban America*. New York, NY: Oxford University Press.
- Krysan, M. (2002). Whites who say they'd flee: Who are they, and why would they leave. *Demography*, *39*, 675–696.
- Krysan, M., & Bader, M. D. M. (2009). Racial blind spots: Black-White-Latino differences in community knowledge. *Social Problems*, *56*, 677–701.
- Krysan, M., & Crowder, K. (2017). *Cycle of segregation: Social processes and residential stratification*. New York, NY: Russell Sage Foundation.
- Kuk, J., Schachter, A., Faber, J. W., & Besbris, M. (2021). The COVID-19 pandemic and the rental market: Evidence from Craigslist. *American Behavioral Scientist*. Advance online publication. <https://doi.org/10.1177/00027642211003149>
- Laidley, T. M. (2014). The privatization of college housing: Poverty, affordability, and the U.S. public university. *Housing Policy Debate*, *24*, 751–768.
- Lareau, A. (2014). Schools, housing, and the reproduction of inequality. In A. Lareau & K. Goyette (Eds.), *Choosing homes, choosing schools* (pp. 169–206). New York, NY: Russell Sage Foundation.
- Lee, B. A., Oropesa, R. S., & Kanan, J. W. (1994). Neighborhood context and residential mobility. *Demography*, *31*, 249–270.
- Logan, J. R., & Alba, R. (1993). Locational returns to human capital: Minority access to suburban community resources. *Demography*, *30*, 243–268.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge, MA: Cambridge University Press.
- Massey, D. S. (2005). Racial discrimination in housing: A moving target. *Social Problems*, *52*, 148–151.
- Massey, D. S., & Denton, N. A. (1993). *American apartheid: Segregation and the making of the underclass*. Cambridge, MA: Harvard University Press.
- McAvay, H. (2018). How durable are ethnoracial segregation and spatial disadvantage? Intergenerational contextual mobility in France. *Demography*, *55*, 1507–1545.
- McLaughlin, R., & Young, C. (2018). Data democratization and spatial heterogeneity in the housing market. In C. Herbert, J. Spader, J. J. Molinsky, & S. Rieger (Eds.), *A shared future: Fostering communities of inclusion in an era of inequality* (pp. 126–139). Cambridge, MA: Harvard Joint Center for Housing Studies.
- Mosteller, F., & Wallace, D. L. (1963). Inference in an authorship problem: A comparative study of discrimination methods applied to the authorship of the disputed Federalist Papers. *Journal of the American Statistical Association*, *58*, 275–309.
- National Multifamily Housing Coalition. (2016). *Quick facts: Resident demographics* [Fact sheet]. Retrieved from <https://www.nmhc.org/research-insight/quick-facts-figures/quick-facts-resident-demographics/>
- Oliveri, R. C. (2010). Discriminatory housing advertisements on-line: Lessons from Craigslist. *Indiana Law Review*, *43*, 1125–1183.
- Pager, D., & Shepherd, H. (2008). The sociology of discrimination: Racial discrimination in employment, housing, credit, and consumer markets. *Annual Review of Sociology*, *34*, 181–209.
- Palm, R., & Danis, M. A. (2001). Residential mobility: The impacts of web-based information on the search process and spatial neighborhood choice patterns. *Urban Geography*, *22*, 641–655.
- Pilkaskas, N., & Michelmore, K. (2019). The effect of the earned income tax credit on housing and living arrangements. *Demography*, *56*, 1303–1326.
- Quillian, L., & Pager, D. (2001). Black neighbors, higher crime? The role of racial stereotypes in evaluations of neighborhood crime. *American Journal of Sociology*, *107*, 717–767.
- Rae, A. (2015). Online housing search and the geography of submarkets. *Housing Studies*, *30*, 453–472.

- Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., . . . Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, *58*, 1064–1082.
- Rosen, E. (2014). Rigging the rules of the game: How landlords geographically sort low-income renters. *City & Community*, *13*, 310–340.
- Rosen, E. (2017). Horizontal immobility: How narratives of neighborhood violence shape housing decisions. *American Sociological Review*, *82*, 270–296.
- Rosenblatt, P., & DeLuca, S. (2012). “We don’t live outside, we live in here”: Neighborhood and residential mobility decisions among low-income families. *City & Community*, *11*, 254–284.
- Sampson, R. J. (2012). *Great American city: Chicago and the enduring neighborhood effect*. Chicago, IL: University of Chicago Press.
- Sampson, R. J., & Raudenbush, S. W. (2004). Seeing disorder: Neighborhood stigma and the social construction of “broken windows.” *Social Psychology Quarterly*, *67*, 319–342.
- Sampson, R. J., & Sharkey, P. (2008). Neighborhood selection and the social reproduction of concentrated racial inequality. *Demography*, *45*, 1–29.
- Schachner, J. N., & Sampson R. J. (2020). Skill-based contextual sorting: How parental cognition and residential mobility produce unequal environments for children. *Demography*, *57*, 675–703.
- Schachter, A., & Besbris, M. (2017). Immigration and neighborhood change: Methodological possibilities for future research. *City & Community*, *16*, 244–251.
- Sharkey, P., & Faber, J. W. (2014). Where, when, why, and for whom do residential contexts matter? Moving away from the dichotomous understanding of neighborhood effects. *Annual Review of Sociology*, *40*, 559–579.
- South, S. J., Crowder, K., & Pais, J. (2011). Metropolitan structure and neighborhood attainment: Exploring intermetropolitan variation in racial residential segregation. *Demography*, *48*, 1263–1292.
- Stiel, J., & Jordan, R. (2018). Household neighborhood decision making and segregation. In C. Herbert, J. Spader, J. Molinsky, & S. Rieger (Eds.), *A shared future: Fostering communities of inclusion in an era of inequality* (pp. 114–124). Cambridge, MA: Harvard Joint Center for Housing Studies.
- Swaroop, S., & Krysan, M. (2011). The determinants of neighborhood satisfaction: Racial proxy revisited. *Demography*, *48*, 1203–1229.
- Taddy, M. (2013). Multinomial inverse regression for text analysis. *Journal of the American Statistical Association*, *108*, 755–770.
- Timberlake, J. M., & Johns-Wolfe, E. (2017). Neighborhood racial composition and gentrification in Chicago and New York 1980 to 2010. *Urban Affairs Review*, *53*, 236–272.
- U.S. Census Bureau. (2017). *2012–2016 American Community Survey 5-year estimates*. Retrieved from https://www.socialexplorer.com/tables/ACS2016_5yr/
- van Ham, M., Boschman, S., & Vogel, M. (2018). Interpreting neighborhood choice in a model of neighborhood effects on income. *Demography*, *55*, 1069–1090.
- Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, *28*(2), 3–28.
- Wang, Q., Phillips, N. E., Small, M. L., & Sampson, R. J. (2018). Urban mobility and neighborhood isolation in America’s 50 largest cities. *Proceedings of the National Academy of Sciences*, *115*, 7735–7740.
- Wood, H. (2014). When only a house makes a home: How home selection matters in the residential mobility decisions of lower-income, inner-city African American families. *Social Service Review*, *88*, 264–294.

Max Besbris

besbris@wisc.edu (corresponding author)

Besbris • Department of Sociology, University of Wisconsin–Madison, Madison, WI, USA; <https://orcid.org/0000-0002-2166-0345>

Schachter • Department of Sociology, Washington University, St. Louis, MO, USA; <https://orcid.org/0000-0002-7404-4140>

Kuk • Department of Political Science, University of Oklahoma, Norman, OK, USA; <https://orcid.org/0000-0003-0772-1110>