NBER WORKING PAPER SERIES

HOUSEHOLD MOBILITY, NETWORKS, AND GENTRIFICATION OF MINORITY NEIGHBORHOODS IN THE US

Fernando V. Ferreira
Jeanna H. Kenney
Benjamin Smith

Working Paper 31480
http://www.nber.org/papers/w31480

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2023

We are grateful for support of the Research Sponsors Program of the Zell/Lurie Real Estate Center. We thank Patrick Kline, Enrico Moretti, Daniel Murphy, Stephen Ross, Bryan Stuart, and seminar participants at Wharton Urban Lunch, UC Berkeley DaveFest 2022, Urban Economics Association Meeting, and Syracuse's John Yinger Conference for helpful comments and suggestions. Finally, we thank the DaveFest 2022 organizers, Janet Currie and Thomas Lemieux, for excellent comments and suggestions that helped shape this research, and also the valuable suggestions of two anonymous referees. We are grateful for support of the Research Sponsors Program of the Zell/Lurie Real Estate Center. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Household Mobility, Networks, and Gentrification of Minority Neighborhoods in the US
Fernando V. Ferreira, Jeanna H. Kenney, and Benjamin Smith
NBER Working Paper No. 31480
July 2023
JEL No. J61,R0

ABSTRACT

We study how recent gentrification shocks impact Black and Hispanic neighborhoods, including where minority households move to after a shock and if the subsequent spatial distribution of households within a labor market area affects segregation. We first report that household moves from a given neighborhood are concentrated to a few destinations. For minority neighborhoods, destinations tend to have similar minority shares but are farther away from downtown. Those mobility patterns are partially explained by neighborhood networks. We then use Bartik-style labor market income shocks to show that gentrification has many effects. In Black neighborhoods, gentrification increases house prices and reduces the share of Black households while increasing the share of White households. For movers from Black neighborhoods, gentrification increases the share of movers going to top 1 and 2 destinations based on neighborhood networks and increases the share of households moving out of the MSA, but does not change the pattern of households moving to neighborhoods with similar Black shares that are farther away from downtown areas. Hispanic neighborhoods have negligible effects from gentrification. Finally, our model reveals that overall labor market area segregation decreases after a gentrification shock because highly Black neighborhoods become less segregated.

Fernando V. Ferreira
Wharton Real Estate
University of Pennsylvania
430 Vance Hall
3733 Spruce Street
Philadelphia, PA 19104-6301
and NBER
fferreir@wharton.upenn.edu

Jeanna H. Kenney
University of Pennsylvania
419E Vance Hall
3733 Spruce Street
Philadelphia, PA 19104-6301
jhkenney@wharton.upenn.edu

Benjamin Smith
University of Pennsylvania
419C Vance Hall
3733 Spruce Street
Philadelphia, PA 19104
Benjamrs@wharton.upenn.edu

1 Introduction

Neighborhood segregation has been a defining feature of American cities for centuries, first promoted by centralized governmental policies in the late 1800s until the mid-1900s, and then driven by decentralized sorting of households over the past five decades (DuBois, 1899; Franklin, 1956; Cutler, Glaeser, and Vigdor, 1999). Such sorting can lead to extreme segregation if tipping points exist where preferences result in White residents completely abandoning neighborhoods when a certain threshold of Black neighbors is met (Schelling, 1971). Card, Mas, and Rothstein (2008) empirically estimate a distribution of tipping points between 5%-20% and show that higher tipping points are related to more tolerant racial views among White residents.[1] The historical influx of Black residents to central neighborhoods of large US cities was also accompanied by suburbanization of White households (Boustan, 2010), but more recently central cities and downtown areas have gone through a revival process (Baum-Snow and Hartley, 2020; Couture and Handbury, 2020, 2023). Such recent inflow of high-income residents could result in increases in house prices (Guerrieri, Hartley, and Hurst, 2013), more redevelopment (Brueckner and Rosenthal, 2009), promote economic gains to low-income neighborhoods (Ellen and O'Regan, 2011), and ultimately the out-migration of local residents (Hwang and Lin, 2016). This process of minority neighborhood change is usually referred to as gentrification, and household displacement is the focus of a large literature in sociology, exemplified by Lees, Slater, and Wyly (2013).

In this paper we study how recent gentrification shocks impact Black and Hispanic neighborhoods, including where minority households move to after a shock, and if the subsequent spatial distribution of households within a labor market area affects segregation. The starting point of the analysis is a new data set that provides address-level moving decisions for a majority of households in the country, which allows us to observe where households move after

[1]Given the prominent role of household preferences and sorting in the shape of segregation in a city, research has primarily focused on estimating heterogeneous preferences for neighborhood characteristics and public goods (Bayer, Ferreira, and McMillan, 2007; Bajari and Kahn, 2005). Other research has focused on how different aspects of housing markets interact with such preferences (Yinger, 1986; Ross and Yinger, 2002; Wong, 2013; Trounstine, 2018).

a gentrification shock and the characteristics of destination neighborhoods. The data set is provided by GrayHair Software, and it is based on proprietary address changes collected from companies (such as financial services, cable, communications, etc.) combined with county and utility public records. The final data contain almost 100 million unique records from 2009 to 2020, with each observation consisting of old and new addresses for each household, names of the individuals moving, the date of move, and indicators for families or individuals.

Three stylized facts arise from this new mobility data. First, moves across neighborhoods within a labor market area are highly concentrated. Focusing on the 50 largest Metropolitan Statistical Areas (MSAs) in the US, on average 49% of neighborhood moves are to only five zip codes, and the number one destination zip code receives 20% of movers. This concentration is similar for both minority and non-minority neighborhoods. For comparison, the average MSA has 172 neighborhoods, implying that a naïve residential choice model would predict that just 0.6% of moves would be to any neighborhood. Even in simulations (detailed in the paper) that use the MSA-level probability distribution of moves across neighborhoods, the actual concentration of moves is substantially higher than the simulated one. Second, moves out of minority neighborhoods are segregated in the sense that zip code destinations tend to also have a large share of minorities. Such segregated moves are more pronounced for top destinations, and for minority households moving from minority neighborhoods. Third, those destinations for minority neighborhoods tend to be farther away from downtowns, a process which is sometimes described as Black suburbanization (Bartik and Mast, 2021).

That households move to neighborhoods where similar households live matches results from the broader migration literature. The influential work of Altonji and Card (1991) and Card (2001) find that immigrant inflows can be partially explained by local features of cities, such as the pre-existing share of immigrants of the same group. Those migrant connections across distant geographies can also be turned into county-level network indices (Stuart and Taylor, 2021) and place of work and place of residence connections (Bayer, Ross, and Topa, 2008). We apply these ideas to estimate a new neighborhood network index that captures social interactions and local knowledge of neighborhoods. Our index is based on the propensity for two households to move to the same destination, comparing this propensity for a pair of households that lives in the same neighborhood with that of a pair that lives in the same reference group but not in the same neighborhood. Empirically this is achieved by comparing mobility patterns of individuals that live within the same Census Public Use Microdata Area (PUMA) as the reference group, using zip codes within these PUMAs as neighborhoods. We estimate these network indices for all neighborhoods in our sample, and find a strong effect only for the top connections. Mobility to the top two connected neighborhoods are, on average, 136% and 29% higher than baseline mobility, respectively.[2]

Using these mobility data and network indices, we proceed to estimate the impact of gentrification on minority neighborhoods from 2010 to 2020.[3] We adapt the empirical model of Guerrieri, Hartley, and Hurst (2013), which uses Bartik shocks as an exogenous source of variation in local incomes, as in Bartik (1991). The intuition is that local income shocks translate into local housing demand shocks due to increases in wages and population inflows (Blanchard and Katz, 1992). In practice, our Bartik shock is created by starting with the 2010 industry composition of residents within each MSA, and then multiplying it by the income growth by industry between 2010 and 2020 for individuals in the entire US (excluding residents from own MSA).

We present our estimates as the effect of a 1-standard deviation increase in the Bartik shock on a neighborhood with an average Black share. We find that gentrification has five main effects. First, house prices increase by 4.4 percentage points (ppts.) over the 10-year period relative to an average price growth of 45%. Second, the share of Black residents is reduced by almost 0.8 ppts. (relative to an average Black share of 11.6%) while the White share increases by a similar 0.8 ppts. These results indicate that, on average,

[2]Next section presents a number of robustness tests that corroborate the validity of the estimated indices and also compares them with other alternative network indices.

[3]We also complement the mobility data with data on house prices, annual American Community Survey records, and Decennial Census data from 2010 to 2020. See Section 2.1 for details.

recent gentrification shocks change minority neighborhoods as expected, with decreases in affordability that result in reductions in the minority composition of neighborhoods. To the best of our knowledge this is the first study to report those gentrification estimates for the last decade.

The subsequent effects are those on movers from Black neighborhoods. We find that gentrification shocks increase the share of households moving to the top 1 and top 2 most-connected neighborhoods as measured by our neighborhood network index by 1.2 ppts. and 1.5 ppts., respectively. Estimates are not significant for top 5 index destinations. The fourth effect is an increase in the share of households moving out of MSA, which increases by 1.4 ppts. Fifth, we do not observe a differential effect on the characteristics of destination neighborhoods, i.e. households are still moving to neighborhoods with similar Black shares and farther away from downtown, suggesting that gentrification is not changing the process of Black suburbanization.

We also test if those effects vary by race of the mover and find that moves to top connections and out-of-MSA have stronger gentrification effects when only focusing on Black households in Black neighborhoods. Interestingly, we find that gentrification does not impact house prices or ethnic and racial shares in Hispanic neighborhoods, and we consequently also find no effects for movers from Hispanic neighborhoods. This could potentially reflect the different demographic trends in the US, with high Hispanic growth rates in many cities expanding Hispanic neighborhoods (Hwang, 2015).[4]

How does the gentrification shock impact segregation within MSAs? On the one hand, gentrified neighborhoods are becoming more diverse because Black neighborhoods are gaining in White share. But on the other hand, some destination neighborhoods may be becoming less diverse, since some movers are concentrating to neighborhoods that have similar concentration of minorities. Empirically, we first show that segregation by MSA - measured according to the dissimilarity index - had small reductions from 2010 to 2020, with the dis-

[4]That Hispanic households and neighborhoods have different outcomes than Black households and neighborhoods is also observed in different settings, such as in health outcomes (Fernandez et al., 2023).

similarity index declining by 2.4 ppts. on average in the 50 largest MSAs. When we use our estimates on neighborhood demographic change to simulate the effect of a one-standard deviation increase in the Bartik income shock for the average top 50 MSA, we find that these gentrification shocks are decreasing Black-White segregation through the integration of Black neighborhoods by White movers.

Our work provides many contributions to the literature. We study the impact of gentrification on minority neighborhoods and minority households using a new high frequency micro data set of household mobility. Other work using mobility data include Ding, Hwang, and Divringi (2016), who use data from the consumer credit panel for one city to study migration patterns of low credit score residents in gentrifying neighborhoods, and Brummet and Reed (2021) who use a selected longitudinal sample of Census respondents with protected identification keys to study the impact of a city-level college-educated demand shock on the out-migration of less educated renters.

Our study describes the consequences of gentrification to segregation and the spatial distribution of households. It complements Couture et al. (2021) who also consider a similar Bartik shock when studying the impact of rising incomes at the top of the distribution on spatial sorting patterns. Moreover, our measurement of where movers migrate is also relevant for estimating welfare, and may also reveal consequences that are not captured by aggregate standard measures of segregation - such as if a neighborhood is far from downtown.[5]

We also provide a new way of measuring social connections across neighborhoods. Other work has provided alternative measures, such as Bailey, Cao, Kuchler, Stroebel, and Wong (2018) who estimate networks and social connectedness at the county-level based on Facebook. For the broader literature on network connectedess and dyadic regressions see Fafchamps and Gubert (2007) and Graham (2020). For a review of the literature on neighborhoods and social networks, see Topa and Zenou (2015). There is also a large literature outside of economics focused on how social capital interacts with gentrification (e.g., Versey, 2018).

[5]See Vigdor et al. (2002) for a review on whether gentrification harms the poor, and Diamond and McQuade (2019) for how welfare is impacted by neighborhood change.

Our study additionally uncovers a new stylized fact that inter-neighborhood mobility is highly concentrated in a small number of destination neighborhoods. That fact could have implications for econometric models of household sorting that have to make assumptions about consideration choice sets; examples include Dingel and Tintelnot (2021) and Ferreira and Wong (2023). It could also have implications for policy. For example, interventions to help families choose where to live, such as programs designed to help families move to opportunity (Bergman et al., 2023; Schwartz et al., 2017), should consider that households are more likely to move to high-opportunity neighborhoods which are also top destinations associated with their current neighborhood. Likewise, if place-based policies have downstream consequences on household mobility then policymakers should consider the top destinations of treated neighborhoods (Busso et al., 2013; Wheeler, 2022). In this regard, our work also has a connection with the literature interested in mobility and economic outcomes; this research has found that mobility can increase economic opportunity (Chetty, Hendren, and Katz, 2016), social capital can affect the decision to move (Kan, 2007; Büchel et al., 2020), and the decision not to move can affect social capital formation (David et al., 2010).

The rest of the paper proceeds as follows. In Section 2 we describe the data and report key stylized facts, including our new neighborhood network index. In Section 3 we show the econometric framework to estimate the impact of gentrification, and report our main results. Section 4 concludes the paper.

2 Neighborhood Mobility

2.1 Data

The mobility data set used in the paper comes from GrayHair Software, and is called "GrayHair New Mover List Services." The data set has proprietary address changes collected from GrayHair clients, such as financial services and cable and communications companies, in addition to county and utility public records. While other internet-data firms may have somewhat similar data sets, those firms are generally not forthcoming or cannot disclose their data sources. The raw data has more than 200 million records from 2009 to 2019, and each observation contains old and new addresses, the name of the individual(s) moving, the effective date of the move, and an indicator for individuals versus families.[6]

We process the raw data by first validating addresses in GrayHair by matching them to the Census Geocoder. This allows us to map moves to the origin and destination Census geographies, such as zip codes, and we drop moves with addresses that cannot be matched (17% of the observations). Second, we eliminate duplicates since GrayHair collects similar moves from different sources (47% of the remaining observations were deleted in this stage).

The final sample used in this paper has 90.5 million observations, and Appendix Table B1 shows the number of moves by year. Appendix Figure A1 shows a map with number of moves by county in 2015, which demonstrates the national coverage of this data. Counties with the largest number of origin moves are Los Angeles with 289,852 moves, Cook with 183,482 moves, and Maricopa with 179,837 moves.

We benchmark the GrayHair data by comparing it to the American Community Survey (ACS), which asks households if they have moved in the past year.[7] Using data from IPUMS (Ruggles et al., 2022), we consider the coverage of movers into different geographic areas at the PUMA level in Figure A2. Panel (a) reports the coverage across deciles of various demographic groups for the year 2015. Average coverage for a given PUMA characteristic is shown by the decile of that same characteristic; for example, coverage for Black PUMAs are shown by the decile of PUMA Black shares. Reassuringly for the purpose of this study, rates are fairly flat among areas with different racial and ethnic compositions. Coverage is, however, correlated with PUMA median income: while GrayHair covers most moves in the highest-income PUMAs, coverage is around 45% in the lowest-income PUMAs. This could

[6] A limitation of the data set is that GrayHair requires individuals to have old and new addresses, so the first American address of a recent immigrant would not appear in the sample. The same omission happens when individuals move abroad.

[7] This is an imperfect measure of coverage, as many movers in GrayHair moved more than once in a year and the ACS cannot pick up these moves. Thus for some cuts of the data, coverage rates may exceed 100%.

occur because families living in low-income PUMAs have lower usage of financial services and cable companies. Panel (b) reports the coverage over time, which oscillates between 40-60%.

For the final sample we also match first and last names to NamePrism (Ye et al., 2017; Ye and Skiena, 2019), a name-based nationality/ethnicity classification tool. Following Diamond, McQuade, and Qian (2019), we then combine the block-based racial probabilities (using 2010 Census data) and the name-based racial probabilities using Bayes' rule to impute the race of movers in the data set. We assign a household as being White, Black, or Hispanic if the probability of any of those is greater than 50%. We will use these racial assignments to test if gentrification estimates for minority neighborhoods have a differential effect by race of the household.

Finally, for our gentrification estimates we complement GrayHair with other data sources. We use Census 2010 and 2020 to calculate race and ethnicity shares in each zip code, as well as the ACS in 2010 and 2019[8] to calculate the Bartik shock to income described in Section 3.1 (Ruggles et al., 2022). We use data from the Federal Housing Finance Agency (Bogin, Doerner, and Larson, 2019) on zip-code level house price indices.

2.2 Stylized Facts

We start by documenting patterns of residential mobility not yet recognized by the literature. First, the most common destination zip code is the origin zip code, with 24% of moves. Second, moves to a different MSA constitute 27% of the moves. As our work focuses on inter-neighborhood mobility within a labor market area, in this section we restrict the sample to inter-zip code moves within the same MSA, although in Section 3 we study the consequence of gentrification on out-of-MSA moves.[9]

[8]Although we refer to these samples as 2010 and 2019, they are actually the 2007-11 and 2015-19 ACS samples from NHGIS (Ruggles et al., 2022).

[9]Within-zip code moves could be due to a variety of causes. Because housing is heterogeneous, households could be re-optimizing and choosing units that now fit their needs better, or household formation and disintegration could be driving this - see Frost (2020).

These inter-neighborhood moves exhibit a high degree of clustering, as reported in Figure 1, which shows the average share of 2009-11 moves to the highest destination neighborhoods.[10] Across all MSAs nearly 30% of neighborhood moves are to a single outside destination zip code. Given that such concentration may be a consequence of small MSAs having a low number of zip codes, the rest of the analysis uses the 50 most populous MSAs. The mean number of zip codes in those MSAs is 172, and the median is 127. In those MSAs, 20% of moves are to only one zip code destination. 49% and 64% are to top 5 and top 10 destination neighborhoods, respectively. As shown in the same figure, this pattern of concentrated moves is also observed in the largest 10 MSAs, where the average number of neighborhoods is 351. The pattern of concentrated moves remains when we only look at the highest quartile of Black neighborhoods, where 16% and 41% of moves are to the top 1 and the top 5 destination neighborhoods, respectively.

To assess whether this concentration of moves is spurious, we constructed two synthetic data sets. The first draws neighborhood destinations from a distribution with an equal probability for every destination neighborhood in the MSA. The second instead uses a distribution equal to the observed MSA-level mobility propensity of moving to every neighborhood; details on this procedure can be found in Appendix C.[11] We repeat this procedure 100 times for four MSAs: San Francisco-Oakland-Berkeley, Philadelphia-Camden-Wilmington, Cleveland-Elyria, and Buffalo-Niagara Falls. Comparisons of actual and simulated moves for those four MSAs are shown in Figure 2. As is clear, the move shares to top zip codes far exceeds the levels exhibited when mobility is equal across zip codes, and even when mobility is assessed using the MSA-level empirical distribution of moves. In Philadelphia, for example, about 16% of the moves are concentrated in the top 1 destination while simulations show similar top shares that are smaller than 3%.

Households tend to move to other neighborhoods which are demographically similar.

[10] We use this earlier period for stylized facts because we also use it for the estimation of network indices in Section 2.3 and then use later periods as outcomes in the estimation of the impact of gentrification in Section 3. Stylized facts are qualitatively similar when using moves from 2009-2020.

[11] We thank Patrick Kline for recommending such approach and providing baseline simulation code.

Panel (a) of Figure 3 shows that among the origin neighborhoods with the highest Black shares, movers tend to go to other neighborhoods with high Black shares. Such correlation is stronger when constrained to the top destination. Panel (b), which considers only movers who are likely Black, shows that the top destination for Black movers who leave Black neighborhoods is to a neighborhood of almost-identical Black share. These figures show that neighborhood mobility tends to be segregated, in the sense that origin and destination neighborhoods have similar minority shares. The pattern for Hispanic movers is similar, and is reported in Figure A3.

Neighborhood mobility over the sample timeframe also exhibits a stark shift in how far away migrants move from the center of the MSA they live in. Figure 4 reports the average distance of the destination neighborhood from the center of a given MSA by origin neighborhood Black share ventiles.[12] For the neighborhoods with the top 5% highest Black share (which, on average, are about two-thirds Black), out-migrants move 2 miles farther away from the center of the city on average. White neighborhoods, on the other hand, show the opposite, with high degrees of mobility to more central neighborhoods.

A stylized example of these patterns of residential mobility is shown in Figure 5, which maps the share of moves from zip code 19123 (hereafter referred to as "Northern Liberties") to all other zip codes in the core of the Philadelphia metropolitan area. Northern Liberties is a typical gentrifying zip code: the White share increased from 41% to 51% from 2010 to 2020, while the Black share fell from 42% to 26%. This example shows, as highlighted earlier, that the vast majority of moves from Northern Liberties are to a small number of zip codes, mapped in dark red in Panel (a). Moreover, those zip codes tend to be more distant from downtown than the origin zip code. However, mobility is different by race as shown in Panel (b): while White movers disproportionately moved to far suburbs and other neighborhoods in the core of the city, Black movers disproportionately moved to inner suburbs within the city and several clustered neighborhoods outside the city, re-enforcing the finding in Figure

[12]We follow Holian and Kahn (2015) and measure the center of the city as the coordinates returned by Google Maps.

3.

2.3 Neighborhood Network Index

The concentration of neighborhood moves to a small number of destinations could be explained by household sorting over prices and neighborhood amenities.[13] However, it could also be that, due to existing residential and experiential segregation, the concentration of moves to a few destinations may in part reflect the network of peers with whom households share social capital. Intuitively, sets of households who are friends or peers may move to the same neighborhood to be located near each other to facilitate social interactions or because they are better able to share local knowledge about neighborhoods.

While it is beyond the scope of this paper to identify all sorting mechanisms driving the concentration of moves, it is interesting to ask if households follow similar mobility patterns after a gentrification shock and if networks have any power in explaining mobility as they have in other settings (Card, 2001; Stuart and Taylor, 2021). The nature of our mobility data explained above allows us to estimate a simple neighborhood network index based on the approach of Bayer, Ross, and Topa (2008) and Stuart and Taylor (2021). Following their econometric model, we estimate the propensity for two households to move to the same destination, comparing this propensity for a pair of households that lives in the same neighborhood with that of a pair that lives in the same reference group but not in the same neighborhood. We estimate

$$D^k_{i,i'} = \alpha^k_g + \sum_{j \in g} \beta^k_j R^j_{i,i'} + \varepsilon^k_{i,i'} \tag{1}$$

where i and i' denote two households who reside in the same reference group g using a dataset of paired moves. $D^k_{i,i'}$ is a dummy variable that is equal to 1 if i and i' moved to the same destination neighborhood k, and 0 otherwise. α^k_g denotes the residential reference

[13]In the standard neighborhood choice framework (Bayer, Ferreira, and McMillan, 2007), households maximize utility by choosing neighborhoods based on heterogeneous preferences for local prices and amenities.

group g fixed effect. $R^j_{i,i'}$ is a dummy variable that is equal to 1 if i and i' resided in the same origin neighborhood j, and 0 otherwise. β^k_j captures the additional probability that a pair of movers from the same origin j both locate in destination k. The OLS estimates $\hat{\beta}^k_j$ are the estimated neighborhood network index values.

The inclusion of the reference group fixed effect is designed to control for any correlation in unobserved attributes among individuals residing in the same neighborhood. In practice we define neighborhoods as zip codes, and the larger reference geography g as the PUMA that the majority of a zip code lives in. On average each PUMA contains between 6 and 7 zip codes. Using zip codes and PUMAs has several benefits. As in Bayer, Ross, and Topa (2008), the reference group is a Census geography where the footprint is continuous and the size reflects the underlying population density. More importantly, those geographies have enough moves that allows us to precisely estimate network connections at a neighborhood level. We estimate Equation 1 separately for each destination zip code k in each MSA leveraging the analytical formulas for a regression using only dummy variables.[14] Finally, we estimate β^k_j using moves from 2009 through 2011, so that we can understand the impact of those connections for changes in outcomes due to gentrification between 2010 and 2020.

Interpreting β^k_j as causal requires two strong assumptions. First, there is no difference across neighborhoods in a given reference group in the value of moving to the destination - this guarantees that moves from zip codes within a larger geography g are suitable control zip codes. We provide a robustness test of this assumption by estimating an extended version of Equation 1 that includes controls for observed characteristics of origin neighborhoods for both

[14] When using the sample of all pairs of individuals from a given g, define

$$\hat{\alpha}^k_g = \frac{\sum_i \sum_{i',j(i)\neq j(i')} D^k_{i,i'}}{\sum_i \sum_{i',j(i)\neq j(i')} 1} \qquad \hat{\gamma}^k_{j_g} = \frac{\sum_i \sum_{i',j(i)=j(i')} D^k_{i,i'}}{\sum_i \sum_{i',j(i)=j(i')} 1}$$

For $\hat{\alpha}^k_g$, the numerator is the number of pairs who both move to k from different j's, and the denominator is the number of pairs from different j's; it is clear that $\hat{\alpha}^k_g = \hat{\alpha}^{k,OLS}_g$. For $\hat{\gamma}^k_j$, the numerator is the number of pairs who both move to k from the same j, and the denominator is the number of pairs from the same j. Then it is clear that $\hat{\beta}^{k,OLS}_j = \hat{\gamma}^k_j - \hat{\alpha}^k_g$. Calculating the regression coefficients this way saves considerable time.

individuals i and i'.[15] Second, household location decisions are not influenced by households from other neighborhoods. Given the stringency of those assumptions, we interpret the index values as informative conditional correlations and we caution interpreting these results as causal. That serves the purposes of this work; in the next section we test if mobility patterns to top-ranked neighborhoods, according to the estimated value of the network index, change after a gentrification shock.

The resulting estimates for top connections are shown in Figure 6. Among the largest 50 MSAs, the highest-connection neighborhood has an estimate of 0.0411, meaning that a pair of movers from the same zip code are 4.11 percentage points more likely to move to a given destination zip code than pairs from different zip codes within the same PUMA. Considering a baseline probability of 2.98% (the probability that pairs from the same PUMA but from different zip codes move to the same destination zip code), this constitutes a 138% increase over the baseline co-mobility probability. For the second-most connected neighborhood, this falls to 0.86 percentage points and a 29% increase in the co-moving probability. The remaining indices are very close to zero.[16]

It is additionally worth noting that while the index is picking up high degrees of co-mobility, it is highly correlated with the move shares: the destination zip code with the largest number of moves from a given origin zip code is the same as the zip code with the highest value of β_j^k approximately 80% of the time. We also tested several other potential

[15]In particular, we estimate

$$D_{i,i'}^k = \alpha_g^k + \sum_g \lambda_g^k |x_j - x_{j'}| + \upsilon_g^k (x_j + x_{j'}) + \sum_{j \in g} \beta_j^k R_{i,i'}^j + \varepsilon_{i,i'}^k \tag{2}$$

where the vector x_j contains the Black share, Hispanic share, median income, median house price, and the share of college educated adults in 2011; this specification controls for intra-PUMA neighborhood differences to assure paired movers are from neighborhoods with similar baseline demographics. This specification follows the dyadic regression literature, e.g. Fafchamps and Gubert (2007).

[16]We tested the statistical significance of those estimates using 200 bootstrapping samples for the Philadelphia MSA, where paired observations are drawn with replacement. We found that 95% of the number one indices are statistical significant, and the same estimate is 94% and 90% for numbers two and three indices, respectively. The literature provides other methods of estimating standard errors in dyadic estimates - see Fafchamps and Gubert (2007) and Graham (2020). In this work we are less interested in the precise statistical significance of a given index, as we will only focus on understanding mobility patterns before and after gentrification for the very top neighborhood indices.

indices. First, we constructed an index equal to the probability of moving to zip code k given origin j minus the unconditional probability of moving to k. Second, we instead subtracted the probability of moving to k conditional on being from origin reference group g. Finally, we estimate a robust version of the index with controls as discussed above. The results do not change substantively: we find that they are relatively similar in magnitude, as can be seen in Appendix Figure A4. Because our main focus in the empirical section is on the rank of neighborhoods according to the index, we also compare the stability of the top-ranked neighborhood according to the index; we find that the most-connected neighborhood in each of the alternative indices is in the top 2 of our preferred specification 87%, 89%, and 70% of the time, respectively.[17]

Despite these positive results, it still could be the case that the network indices reflect spurious correlations in the data. For instance, because we are reporting the maximum of a large number of draws from many random variables, these large effects among highly-connected neighborhoods could be due to chance. To assess this possibility, we turn to the simulated MSA-level data sets described in Section 2.2, which we use to calculate our index 100 times per MSA. The results for the four selected MSAs are shown in Figure 7, which indicate that the most-connected neighborhoods exhibit higher co-mobility than would be expected under common moving probabilities. This is always true for the most connected neighborhood, and the effect decays by the third-most connected neighborhood in all MSAs.

[17] We thank an anonymous referee for this suggestion. Stuart and Taylor (2021) propose an alternative measure of the direct additional probability that pairs of movers from a particular zip code move to the same destination zip code. Defining $P_{g,k} = E[D^k_{i,g(i)}]$, the probability that a mover from g moves to k, and $\mu_{j,k} = E[D^k_{i,j(i)}|D^k_{i',j(i')} = 1]$, the probability that a mover from j goes to k conditional on another individual from the same j moving to k. The additional probability of moving to a given neighborhood k relative to the baseline could then be $\mu^k_j - P^k_g$. However, Stuart and Taylor (2021) show that our estimator $\beta^k_j = P^k_g(\mu^k_j - P^k_g)$, meaning the direct effect is being scaled by the baseline probability of moving to a given neighborhood; thus, for popular neighborhoods, we might be over-estimating the impact of social networks on moving decisions in our setting. While we suspect this is not the case as zip codes tend to be more similar in size than in other empirical applications, we estimate the baseline probabilities $\hat{P}_{g,k}$ as in Stuart and Taylor (2021) and divide our estimated network index by them to test the importance of this. We found that the most-connected neighborhood in this alternative index was in our preferred specification's top 2 destinations 90% of the time.

3 Gentrification of Minority Neighborhoods

3.1 Model

We exploit Bartik-style labor income shocks to estimate the impact of gentrification on minority neighborhoods. We estimate these plausibly exogenous shocks to local housing demand by measuring variation in the national earnings by industry between 2010 and 2019.[18] To construct the MSA-level Bartik shocks, we begin by calculating the industry composition in each MSA j in 2010 using the ACS sample, where industries are represented by two-digit NAICS codes. We only consider employed individuals aged 25-55. We then calculate the income growth for each industry from 2010-2019 (using the 2019 ACS sample), across every MSA in the US *except* j. We apply these industry-level leave-one-out income growth numbers to the industry composition in j in 2010 to obtain a plausibly exogenous measure for expected income growth from 2010 to 2019 in that MSA over that time period.

We model outcomes as in Guerrieri, Hartley, and Hurst (2013) using the following equation:

$$Y_{jm} = \mu_m + \theta_B(\hat{I}_m \times B_{jm}^{2010}) + \omega_B B_{jm}^{2010} + \theta_H(\hat{I}_m \times H_{jm}^{2010}) + \omega_H H_{jm}^{2010} + \epsilon_{jm} \tag{3}$$

In this setting, Y_{jm} is an outcome of origin zip code j in metropolitan area m. B_{jm}^{2010} is the Black share in the neighborhood in 2010, measured using the Census, while H_{jm}^{2010} is the Hispanic share in the neighborhood in 2010. Our coefficients of interest are θ_B and θ_H, which captures the effect of the Bartik income shock ($\hat{I}_m$) from 2010-2019 interacted with Black or Hispanic share, respectively. μ_m is an MSA fixed effect, and we report standard errors clustered at the MSA level.

Descriptive statistics for the variables used in the estimation are displayed in Table 1. Our preferred sample uses zip codes in the top 50 MSAs by population, which leaves us with

[18]See Bartik (1991) and Blanchard and Katz (1992) for how Bartik shocks result in migration from other areas and increases in house prices.

a final sample 8,576 origin neighborhoods.[19] The top panel displays summary statistics for our dependent variables, while the bottom describes our explanatory variables. Of note, the average zip code in the biggest 50 MSAs saw a decrease in White share over the ten year period from 2010-2020, a smaller decrease in Black share, and an increase in the Hispanic share. The average zip code in our sample is located in an MSA that experienced a Bartik income shock of 18.5%.

3.2 Aggregate Impacts on Origin Neighborhoods

We first consider outcomes for origin minority neighborhoods. We estimate the effect of a gentrification shock on house prices, as measured by the percent change in FHFA zip code-level house price indices from 2010-2020, and on the change in Black, Hispanic, and White share, as measured by the difference in the race share of the zip code population from the 2010 to 2020 Census. Results are displayed in Table 2.

The results in Column 1 highlight that Black neighborhoods on average experienced a smaller increase in nominal house prices from 2010 to 2020. The Share Black coefficient is -1.278, meaning that a neighborhood with an average Black share of 11% would have a price growth rate that is fourteen percentage points smaller than the average of 44.5%. However, the interaction of the Bartik shock with Share Black shows that gentrification partially closed that gap. In the rows "Treatment Effect (Black)" and "Treatment Effect (Hispanic)," we present these estimates as the effect of a 1-standard deviation increase in the Bartik shock (0.041) on a neighborhood with average Black shares (11%) or average Hispanic shares (13%). The average share Black zip code experiencing a one standard deviation increase in income shock had a treatment effect of 4.0%, which is about 10% of the total change in prices during that time period.

Columns 2, 3, and 4 report results for the Black, Hispanic, and White share change, respectively. Black neighborhoods on average saw small increases in Black shares, negligible

[19] Due to data limitations, the FHFA does not provide house price indices for all zip codes, and thus that specification has fewer observations.

decreases in Hispanic shares, and decreases in White shares during the 2010s. But gentrification changes this direction, with Black neighborhoods experiencing a larger income shock having a decrease in the Black share and an increase in the White share. An average Black share zip code with a one standard deviation increase in income shock has a 0.8 percentage point decrease in Black share, which is nearly eight times the mean Black share change throughout our sample of zip codes in top 50 MSAs (the mean Black share change was quite small at 0.1 ppts.). Conversely, that same zip code would see a 0.7 percentage point *increase* in the White share. Notably, while Hispanic neighborhoods see an increase in the share Hispanic on average over this time period, there is no significant effect of a Bartik shock on any of the demographic shares in these neighborhoods.

Because racial shares may be correlated with other origin zip code characteristics, we additionally re-specify equation 3 by controlling for the natural log of median zip code income and the share of residents in the zip code who lived in a different residence a year prior according to the ACS. The results are displayed in Table 3. Neither of these controls when interacted with the Bartik shock has a significant effect on the outcomes. Further, the effect of the Bartik shock interacted with the Black share remains similar across all specifications. This further suggests that our results are indeed being driven by the racial share of the origin zip codes.

As another robustness check, we re-specify our model to exclude MSA fixed effects and include the Bartik income shock as an explanatory variable. This allows us to leverage across-MSA variation in Black share to identify our key parameter, and also test if the interactions results are not driven by the independent effect of the Bartik income shock. The results are shown in Appendix Table B2 and are both quantitatively and qualitatively in line with the results above. We additionally estimate a model using an indicator for the quartile of zip code Black share and Hispanic share as opposed to our continuous measures of minority shares - see Appendix Table B3. Results remain stronger for the Black and White share change, and the effects are concentrated amongst zip codes in the highest Black share quartiles.

In summary, Black zip codes experiencing a larger city-wide labor income shock end up having higher house prices, lower Black shares, and higher White shares over time. Negligible effects are found for Hispanic neighborhoods. However, it remains unclear how gentrification impacted households moving out of gentrified neighborhoods. Thus we now turn to our results regarding where movers from the origin zip codes go.

3.3 Impact on Movers

We have shown via stylized evidence that inter-neighborhood mobility is highly concentrated to a few destination neighborhoods, moves tend to happen between neighborhoods that are demographically similar, and this correlation of origin-destination characteristics is stronger for top destination neighborhoods. We now turn to estimating whether these patterns remain the same or change after a gentrification shock.

Therefore, the second set of outcomes we consider is for those households moving out of minority neighborhoods. We estimate the effect of a gentrification shock on the share of moves to the top 1, to either top 1 or 2 (labeled top 2), and all top 5 destinations as measured by the neighborhood network index described in Section 2.3. We consider the share of moves out of that zip code but within the MSA from 2012-2019 (as our index is estimated with moves from 2009-2011). We also consider the share all moves out of the zip code that are to a destination outside of the MSA, again using moves from 2012-2019. Results are displayed in Table 4.

The first three columns display results for the share of moves to the top network indices. High Black share neighborhoods tend to send fewer movers to these destinations on average. However, higher Black share neighborhoods which also see a higher income shock send more movers to top network destinations. An average Black share zip code with a one standard deviation income shock send an extra 1.1 and 1.3 percentage points more movers to their top 1 and either top 2 network destinations, respectively. Results are insignificant when considering moves to all top 5 destinations, suggesting it is the top two connections which matter most for neighborhood movement. In Column 4, we consider the share of movers who leave the MSA entirely, seeing that again on average, higher Black share neighborhoods send fewer movers out of the MSA, but those which experience a gentrification shock send more to out-of-MSA destinations. This amounts to a 1.8 percentage point treatment effect. There is no significant effect of an income shock in Hispanic neighborhoods on the types of destinations movers go to. Table 5 also shows that estimates are not sensitive to including other controls for income and share of movers. The results above are robust to a number of tests.[20]

Thus far, we have provided evidence that income-driven gentrification shocks impact Black neighborhoods in two distinct ways. Origin neighborhoods experience higher home prices, decreased Black share of residents, and increased White share. Those who move out of the neighborhood are more likely to go to the top connections and out of the MSA. We have additionally provided evidence that there are no significant effects of a labor income shock on higher Hispanic share neighborhoods for both aggregate outcomes and the outcomes for movers.

3.4 Characteristics of Destinations

We now turn to considering the effects of gentrification shocks on the characteristics of the destinations movers go to. In Table 6, we consider the Black share of the destinations relative to the Black share of the origin zip code (Columns 1 and 2), and the distance to the MSA center of the destinations relative to the origin (Columns 3 and 4). For each mover out of an origin zip code j, we calculate the difference in Black share of the destination zip code k and origin zip code. Thus, positive numbers indicate that destination zip codes have a

[20] We again re-estimate the model without MSA fixed effects (see Appendix Table B4) and by using indicators for the quartile of Black share and Hispanic share instead of the continuous measure of Black share (see Appendix Table B5). Similarly, results are qualitatively robust to removing the MSA-level controls and controlling for the Bartik income shock. Additionally, when controlling for quartile race shares as opposed to continuous measures, results are again strongest for the share of movers going to the top 1 and top 2 destinations, and are concentrated amongst zip codes in the fourth quartile of Black Share.

higher Black share than the origin. In Column 1, we average these differences (i.e., relative Black shares) over all movers. In Column 2, we consider only movers to the top destination zip code as defined by our network index. Note that the averages for these variables are quite small, which is a consequence of the pattern of segregated moves described in Figure 3. Across both measures, we find no statistically significant evidence that income shocks resulted in movers going to neighborhoods with different racial compositions.

In columns 3 and 4, we calculate the difference in distance to the MSA center of the destination zip code k and origin zip code averaged across all movers (3) and for only movers to the top destination (4). Positive numbers imply moving further away from the downtown. The average of the relative distance variable in Columns 3 and 4 is negative because most movers are White and they tend to move closer to downtown. The Share Black coefficient in Columns 3 and 4 is positive, indicating (as in Figure 4) that households from Black neighborhoods tended to suburbanize during that time period.[21] Across all of the destinations of movers, a Bartik income shock mitigates this effect; however, this is not the case for the top network destination alone.

3.5 Heterogeneity by Household Race

It is possible that certain types of movers experience gentrification shocks differently. Thus far we have only considered a heterogeneous response to gentrification by the Black share of the origin zip code. We now also consider whether Black *movers* experience the shock differently.

In Table 7 we replicate the results of Table 4, considering again where movers go in the face of a gentrification shock. Now, however, we consider only Black movers.[22] The effect sizes are nearly doubled when we consider only Black movers, and more neighborhood

[21] These results qualitatively match the Black suburbanization patterns of Bartik and Mast (2021) and that destination neighborhoods are similar with respect to other observables as in Brummet and Reed (2021).

[22] A mover is again classified as Black based on our imputation procedure using the mover's name and census block, as described in Section 2.1. We assume a mover is Black if their probability of being Black is greater than 50%.

connections matter as there is now a significant effect of sending people to any of the top 5 destinations. An average Black share zip code with a one standard deviation higher income shock will send 2 percentage points more Black movers to the top 1 destination, which is 13% of the mean share of Black movers to the top destination across zip codes.

We also investigate whether our results regarding the characteristics of destinations change when we consider Black movers only. Column 1 of Table 8 replicates results of Column 1 of Table 6, asking whether high Black share neighborhoods experiencing a gentrification shock send movers to neighborhoods that have an even higher Black share, again this time only considering Black movers. We find no statistically significant evidence of this occurring. Column 2 of Table 8 replicates Column 3 of Table 6. Similarly, average relative distance to downtown decreases for Black movers after a gentrification shock, though to a lesser extent than when we considered all movers. An average Black share zip code with a one standard deviation higher income shock will send Black movers to destinations 0.636 miles closer to the MSA center than their origin, which is 36% of the mean relative distance to downtown for Black movers.

3.6 Aggregate Impacts on Segregation

We have provided evidence that gentrification has impacts on both origin Black neighborhoods and the types of destination neighborhoods to which movers go. We now ask whether these effects lead to an impact on aggregate segregation measures. Researchers have developed many metrics of segregation, but for clarity we focus on the dissimilarity index,[23] a measure of evenness of racial/ethnic groups in the city. The Black-White index reports the percent of Black residents who would have to move so that all zip codes would have the same share of Black residents. We also report an analogous Hispanic-White dissimilarity index.

First, we simply calculate the Black-White dissimilarity index in 2010 and 2020 and

[23] We implement this using the package `seg` from Reardon and Firebaugh (2002), which also reports on various other measures of segregation. We additionally tested our results using the Theil index and the Isolation index and found the same results, which are available upon request. For all indices, we use zip codes as the smaller geographic unit inside the MSA.

report this in Table 9. Segregation is a long-standing pattern in American cities, and it did not fall substantially over the past decade; the dissimilarity index decreased on average by 2.4 ppts. (4.4%) over the last decade for the top 50 MSAs, as can be seen in Figure 8. For comparison, Cutler, Glaeser, and Vigdor (1999) report declines of 16.7 ppts. or 23% in the average dissimilarity index from 1970 to 1990, a period known for large declines in Black-White segregation.

To understand the effect that gentrification has on residential segregation, we predict the changes in racial group shares across zip codes using the estimates from Equation 3. The first simulation uses the true MSA-level shock, while the second simulation uses the true MSA-level shock plus one standard deviation.[24] Table 8 reports the results of these partial equilibrium simulations. As expected, using our estimates to construct a simulated dissimilarity index in 2020 closely approximates the true dissimilarity index in 2020. The numbers are 0.511 for the true Black-White index and 0.509 for the simulated Black-White index, and 0.387 for the true Hispanic-White index and 0.365 for the simulated Hispanic-White index.

When we increase the Bartik shock by one standard deviation for all MSAs and recalculate their dissimilarity indices, we find that our estimates imply the average top 50 MSA's Black-White segregation decreases by 1.4 ppts. and Hispanic-White segregation decreases by 0.7 ppts. These simulated changes in Bartik shocks, however, are quite large as they represent the standard deviation for all MSAs rather than just the top 50 which have a smaller variance. When we use the 1 SD income shock for the top 50 MSAs, these numbers fall to 0.4 ppts. and 0.2 ppts., respectively. These suggest that a large MSA receiving a 1 SD larger income shock would have experienced a 15% and 5% larger decline in Black and Hispanic segregation, respectively, in the 2010s relative to the true decline.

Given the structure of our estimate, the reduction in segregation after a gentrification

[24]We do these simulations by using the coefficients in Table 2 to estimate the change in the Black, White, and Hispanic share. We add the changes to the 2010 share, trim the shares to be between 0 and 1, and multiply the shares by the 2020 zip code population. We then calculate the dissimilarity index using these populations.

shock must be driven by the effects of integration in Black and Hispanic neighborhoods. This can be seen by considering the extreme example of a 100% White neighborhood; our estimates of changes in Black, White, and Hispanic shares for this neighborhood would be unaffected by any of the four main estimates in Table 2 because the share Black and Hispanic share are 0%. This suggests that large gentrification shocks can have positive effects on the patterns of segregation across cities, in contrast to the results in Bartik and Mast (2021).

4 Conclusion

The revival of American cities and their central neighborhoods since the 1980s resulted in the gentrification of many neighborhoods. A large literature in economics, sociology, and urban planning followed, concerned not only with overall neighborhood changes but also with the fate of households displaced by gentrification.

This paper contributes to this literature in several meaningful ways. First, using a new data set with address-level household moves, we detail a key set of stylized facts about neighborhood mobility: moves from a given origin neighborhood are highly concentrated to a few destinations, they tend to be segregated in the sense of origin-destination neighborhoods having similar minority shares, and there is a trend of minority households moving away from downtowns. Second, we leverage recent gentrification shocks on minority neighborhoods to show that gentrification has resulted in house price increases and reductions in the share of minority households, and uncover that the pre-existing neighborhood mobility patterns within a city and out of a city help determine the mobility patterns induced by displacement. That Black movers are following historical migration patterns in response to local income shocks fits in with the literature on both historical Black and international migration as in Boustan (2010) and Altonji and Card (1991).

Third, we did not find evidence that gentrification has changed the process of minority households moving to segregated neighborhoods away from downtowns. Finally, the net effect of those movements on overall MSA segregation based on the dissimilarity index is to decrease segregation by increasing the White share in Black and Hispanic neighborhoods. This indicates that such measures of segregation may miss important neighborhood dynamics, such as the difference between integration by Black residents moving into White neighborhoods versus White residents moving into Black neighborhoods, and spatial components like how far from downtowns different types of households live.

The finding that pre-existing mobility patterns are reinforced through networks after a gentrification shock points to certain policy implications. For example, interventions to help families choose where to live, such as programs designed to help families move to opportunity (Bergman et al., 2023; Schwartz et al., 2017), should consider that households are more likely to move to high-opportunity neighborhoods which are also top destinations associated with their current neighborhood. Likewise, if place-based policies have downstream consequences on household mobility then policymakers should consider the top destinations of treated neighborhoods (Busso et al., 2013; Wheeler, 2022). Finally, knowing mobility patterns from gentrifying neighborhoods can help policymakers target those who are displaced.

References

Altonji, Joseph G., and David Card. 1991. *Immigration, Trade, and the Labor Market.* Chap. The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives 201–234, University of Chicago Press.

Bailey, Michael, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. 2018. "Social Connectedness: Measurement, Determinants, and Effects." *Journal of Economic Perspectives* 32 (3): 259–80. 10.1257/jep.32.3.259.

Bajari, Patrick, and Matthew E Kahn. 2005. "Estimating Housing Demand With an Application to Explaining Racial Segregation in Cities." *Journal of Business & Economic Statistics* 23 (1): 20–33. 10.1198/073500104000000334.

Bartik, Alexander, and Evan Mast. 2021. "Black Suburbanization: Causes and Consequences of a Transformation of American Cities." Upjohn Institute Working Papers, November.

Bartik, Timothy J. 1991. *Who Benefits from State and Local Economic Development Policies?.* Upjohn Institute for Employment Research.

Baum-Snow, Nathaniel, and Daniel Hartley. 2020. "Accounting for central neighborhood change, 1980–2010." *Journal of Urban Economics* 117 103228. 10.1016/j.jue.2019.103228.

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. 10.1086/522381.

Bayer, Patrick, Stephen L. Ross, and Giorgio Topa. 2008. "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes." *Journal of Political Economy* 116 (6): 1150–1196. 10.1086/595975.

Bergman, Peter, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence F Katz, and Christopher Palmer. 2023. "Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice." Working Paper 26164, National Bureau of Economic Research. 10.3386/w26164.

Blanchard, Olivier, and Lawrence Katz. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity* 23 (1): 1–76.

Bogin, Alexander, William Doerner, and William Larson. 2019. "Local House Price Dynamics: New Indices and Stylized Facts." *Real Estate Economics* 47 (2): 365–398. 10.1111/1540-6229.12233.

Boustan, Leah Platt. 2010. "Was Postwar Suburbanization "White Flight"? Evidence from the Black Migration." *The Quarterly Journal of Economics* 125 (1): 417–443. 10.1162/qjec.2010.125.1.417.

Brueckner, Jan K., and Stuart S. Rosenthal. 2009. "Gentrification and Neighborhood Housing Cycles: Will America's Future Downtowns Be Rich?." *The Review of Economics and Statistics* 91 (4): 725–743. 10.1162/rest.91.4.725.

Brummet, Quentin, and Davin Reed. 2021. "The Effects of Gentrification on Incumbent Residents." Working Paper.

Busso, Matias, Jesse Gregory, and Patrick Kline. 2013. "Assessing the Incidence and Efficiency of a Prominent Place Based Policy." *American Economic Review* 103 (2): 897–947. 10.1257/aer.103.2.897.

Büchel, Konstantin, Maximilian V. Ehrlich, Diego Puga, and Elisabet Viladecans-Marsal. 2020. "Calling from the outside: The role of networks in residential mobility." *Journal of Urban Economics* 119 103277. 10.1016/j.jue.2020.103277.

Card, David. 2001. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics* 19 (1): 22–64. 10.1086/209979.

Card, David, Alexandre Mas, and Jesse Rothstein. 2008. "Tipping and the Dynamics of Segregation*." *The Quarterly Journal of Economics* 123 (1): 177–218. 10.1162/qjec.2008.123.1.177.

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz. 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review* 106 (4): 855–902. 10.1257/aer.20150572.

Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst. 2021. "Income Growth and the Distributional Effects of Urban Spatial Sorting." Working Paper.

Couture, Victor, and Jessie Handbury. 2020. "Urban revival in America." *Journal of Urban Economics* 119 103267. 10.1016/j.jue.2020.103267.

Couture, Victor, and Jessie Handbury. 2023. "Neighborhood Change, Gentrification, and the Urbanization of College Graduates." *Journal of Economic Perspectives* 37 (2): 29–52. 10.1257/jep.37.2.29.

Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor. 1999. "The Rise and Decline of the American Ghetto." *Journal of Political Economy* 107 (3): 455–506. 10.1086/250069.

David, Quentin, Alexandre Janiak, and Etienne Wasmer. 2010. "Local social capital and geographical mobility." *Journal of Urban Economics* 68 (2): 191–204. 10.1016/j.jue.2010.04.003.

Diamond, Rebecca, and Tim McQuade. 2019. "Who Wants Affordable Housing in Their Backyard? An Equilibrium Analysis of Low-Income Property Development." *Journal of Political Economy* 127 (3): 1063–1117. 10.1086/701354.

Diamond, Rebecca, Tim McQuade, and Franklin Qian. 2019. "The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco." *American Economic Review* 109 (9): 3365–94. 10.1257/aer.20181289.

Ding, Lei, Jackelyn Hwang, and Eileen Divringi. 2016. "Gentrification and residential mobility in Philadelphia." *Regional Science and Urban Economics* 61 38–51. 10.1016/j.regsciurbeco.2016.09.004.

Dingel, Jonathan I., and Felix Tintelnot. 2021. "Spatial Economics for Granular Settings." Working Paper.

DuBois, W. E. B. 1899. *The Philadelphia Negro: A Social Study*. Philadelphia, PA: University of Pennsylvania Press.

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. 10.1016/j.regsciurbeco.2010.12.005.

Fafchamps, Marcel, and Flore Gubert. 2007. "The formation of risk sharing networks." *Journal of Development Economics* 83 (2): 326–350. 10.1016/j.jdeveco.2006.05.005.

Fernandez, José, Mónica García-Pérez, and Sandra Orozco-Aleman. 2023. "Unraveling the Hispanic Health Paradox." *Journal of Economic Perspectives* 37 (1): 145–68. 10.1257/jep.37.1.145.

Ferreira, Fernando, and Maisy Wong. 2023. "Estimating Preferences for Neighborhood Amenities Under Imperfect Information." Working Paper.

Franklin, John Hope. 1956. "History of Racial Segregation in the United States." *The ANNALS of the American Academy of Political and Social Science* 304 (1): 1–9. 10.1177/000271625630400103.

Frost, Riordan. 2020. "Are Americans Stuck In Place? Declining Residential Mobility in the US." Joint Center for Housing Studies Research Brief.

Graham, Bryan S. 2020. "Chapter 2 - Dyadic regression." In *The Econometric Analysis of Network Data*, edited by Graham, Bryan, and Áureo de Paula 23–40, Academic Press, . 10.1016/B978-0-12-811771-2.00008-0.

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst. 2013. "Endogenous Gentrification and Housing Price Dynamics." *Journal of Public Economics* 100 45–60. 10.1016/j.jpubeco.2013.02.001.

Holian, Matthew J., and Matthew E. Kahn. 2015. "Household Carbon Emissions from Driving and Center City Quality of Life." *Ecological Economics* 116 362–368. 10.1016/j.ecolecon.2015.05.012.

Hwang, Jackelyn. 2015. "Pioneers of Gentrification: Transformation in Global Neighborhoods in Urban America in the Late Twentieth Century." *Demography* 53 (1): 189–213. 10.1007/s13524-015-0448-4.

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

Kan, Kamhon. 2007. "Residential mobility and social capital." *Journal of Urban Economics* 61 (3): 436–457. 10.1016/j.jue.2006.07.005.

Lees, L., T. Slater, and E. Wyly. 2013. *Gentrification.* Taylor & Francis.

Reardon, Sean F., and Glenn Firebaugh. 2002. "2. Measures of Multigroup Segregation." *Sociological Methodology* 32 (1): 33–67. 10.1111/1467-9531.00110.

Ross, Stephen, and John Yinger. 2002. *The Color of Credit: Mortgage Discrimination, Research Methodology, and Fair-Lending Enforcement.* The MIT Press.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Megan Schouweiler, and Matthew Sobek. 2022. "IPUMS USA: Version 12.0 [dataset]." 10.18128/D010.V12.0, Minneapolis, MN: IPUMS.

Schelling, Thomas C. 1971. "Dynamic models of segregation." *The Journal of Mathematical Sociology* 1 (2): 143–186. 10.1080/0022250X.1971.9989794.

Schwartz, Heather L., Kata Mihaly, and Breann Gala. 2017. "Encouraging Residential Moves to Opportunity Neighborhoods: An Experiment Testing Incentives Offered to Housing Voucher Recipients." *Housing Policy Debate* 27 (2): 230–260. 10.1080/10511482.2016.1212247.

Stuart, Bryan A., and Evan J. Taylor. 2021. "Migration Networks and Location Decisions: Evidence from US Mass Migration." *American Economic Journal: Applied Economics* 13 (3): 134–75. 10.1257/app.20180294.

Topa, Giorgio, and Yves Zenou. 2015. "Chapter 9 - Neighborhood and Network Effects." In *Handbook of Regional and Urban Economics*, edited by Duranton, Gilles, J. Vernon Henderson, and William C. Strange Volume 5. of Handbook of Regional and Urban Economics 561–624, Elsevier, . 10.1016/B978-0-444-59517-1.00009-X.

Trounstine, J. 2018. *Segregation by Design: Local Politics and Inequality in American Cities.* Cambridge University Press.

Versey, H. Shellae. 2018. "A tale of two Harlems: Gentrification, social capital, and implications for aging in place." *Social Science & Medicine* 214 1–11. 10.1016/j.socscimed.2018.07.024.

Vigdor, Jacob L., Douglas S. Massey, and Alice M. Rivlin. 2002. "Does Gentrification Harm the Poor?" *Brookings-Wharton Papers on Urban Affairs* 133–182.

Wheeler, Harrison. 2022. "Locally Optimal Place-Based Policies: Evidence from Opportunity Zones." Working Paper, November.

Wong, Maisy. 2013. "Estimating Ethnic Preferences Using Ethnic Housing Quotas in Singapore." *The Review of Economic Studies* 80 (3): 1178–1214. 10.1093/restud/rdt002.

Ye, Junting, Shuchu Han, Yifan Hu, Baris Coskun, Meizhu Liu, Hong Qin, and Steven Skiena. 2017. "Nationality Classification Using Name Embeddings." *CoRR* abs/1708.07903, http://arxiv.org/abs/1708.07903.

Ye, Junting, and Steven Skiena. 2019. "The Secret Lives of Names? Name Embeddings from Social Media." *CoRR* abs/1905.04799, http://arxiv.org/abs/1905.04799.

Yinger, John. 1986. "Measuring Racial Discrimination with Fair Housing Audits: Caught in the Act." *The American Economic Review* 76 (5): 881–893.

5 Figures

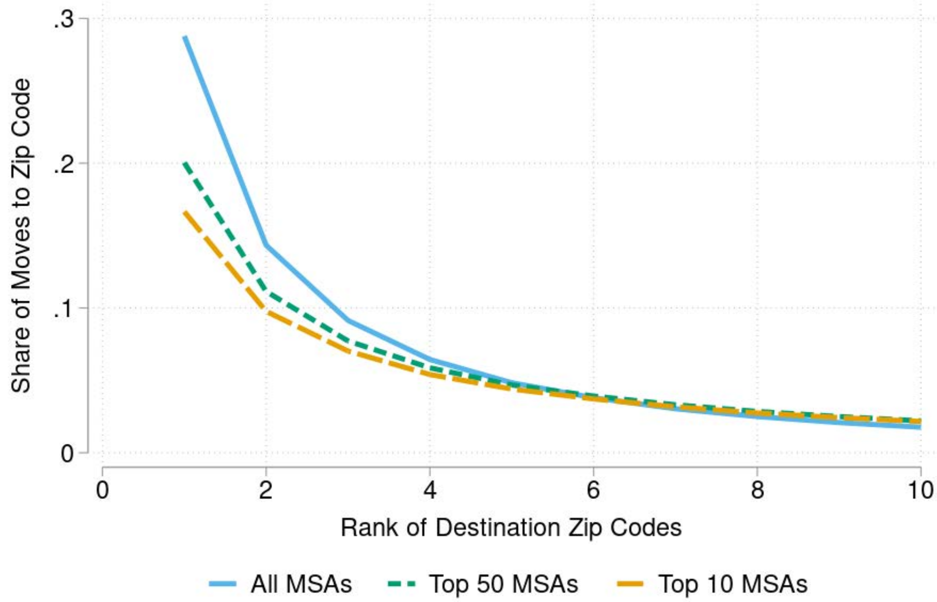
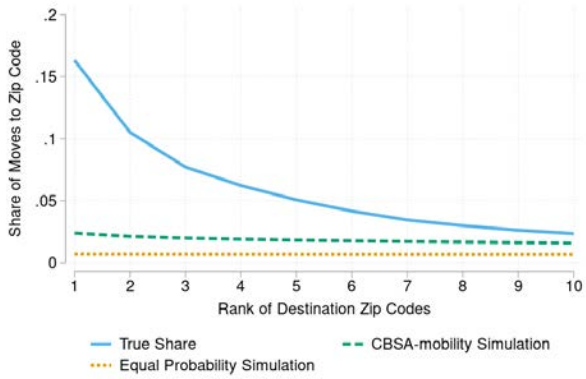
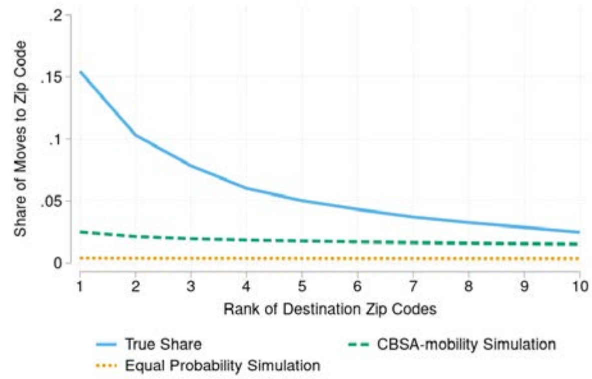


Figure 1: Mobility Concentration

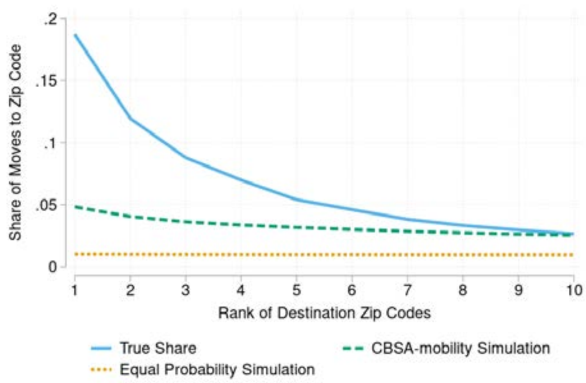
Note: Figure reports the average share of intra-MSA moves out of a zip for its ten highest-migration destinations.



(a) San Francisco-Oakland-Berkeley



(b) Philadelphia-Camden-Wilmington

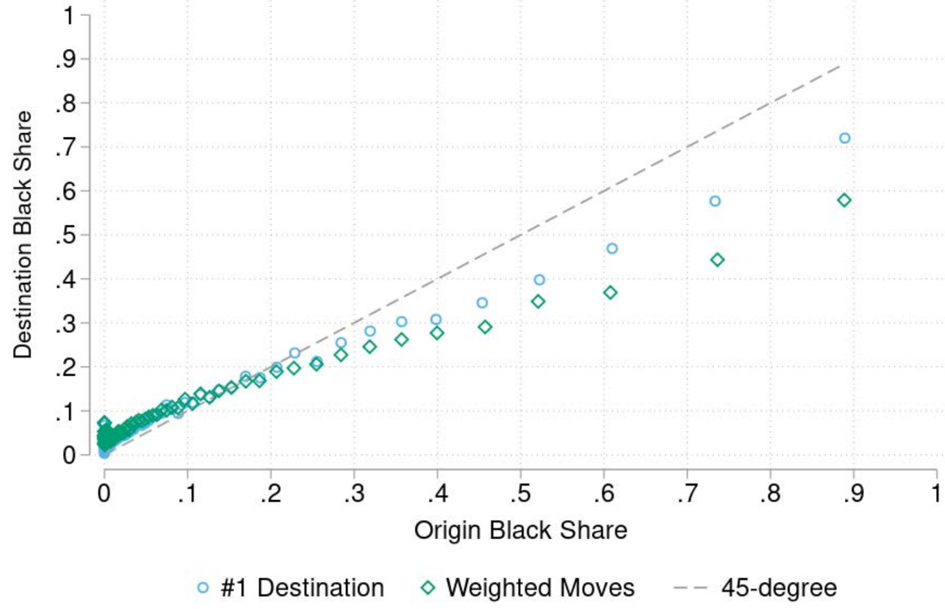


(c) Cleveland-Elyria

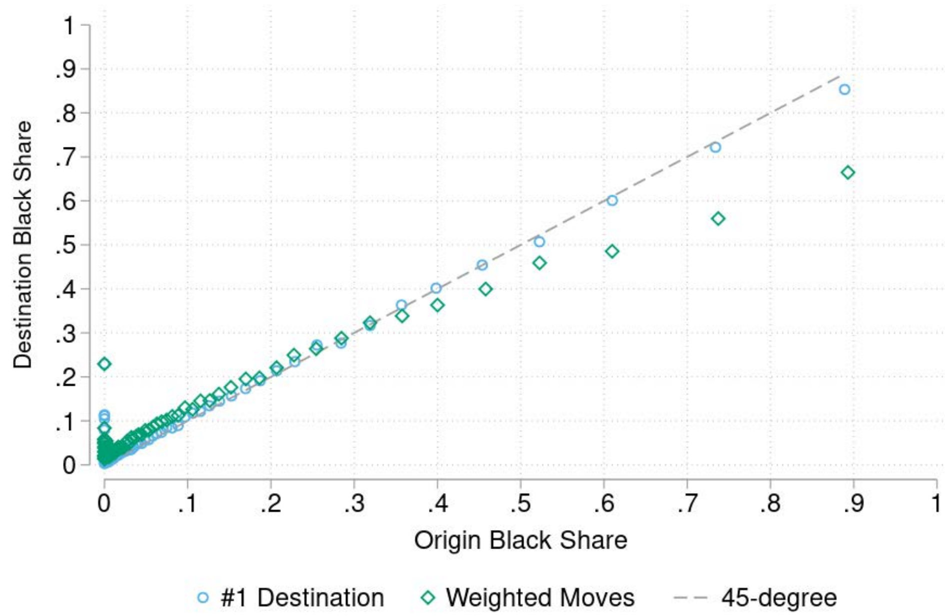
(d) Buffalo-Niagara Falls

Figure 2: Mobility Concentration by MSA

Note: Figure reports the average share of intra-MSA moves out of a zip for its ten highest-migration destinations for four CBSAs. The CBSA-mobility simulation and equal probability simulation reports the same data using the 100 data sets simulated using overall mobility probabilities in the MSA and equal mobility probabilities, respectively. See text for details.



(a) Mobility from Black Neighborhoods, All Movers



(b) Mobility from Black Neighborhoods, Black Movers

Figure 3: Mobility Concentration out of Black Neighborhoods

Note: Panel (a) reports the Black share of the number one and the average Black share of destination neighborhoods by origin neighborhood Black percentile. Panel (b) reports the same but for movers who are Black. The regression estimates of the destination Black share on the origin Black share for #1 destinations and weighted moves are 0.88 and 0.54 in Panel (a) and 0.89 and 0.64 in Panel (b), respectively. See text for details.

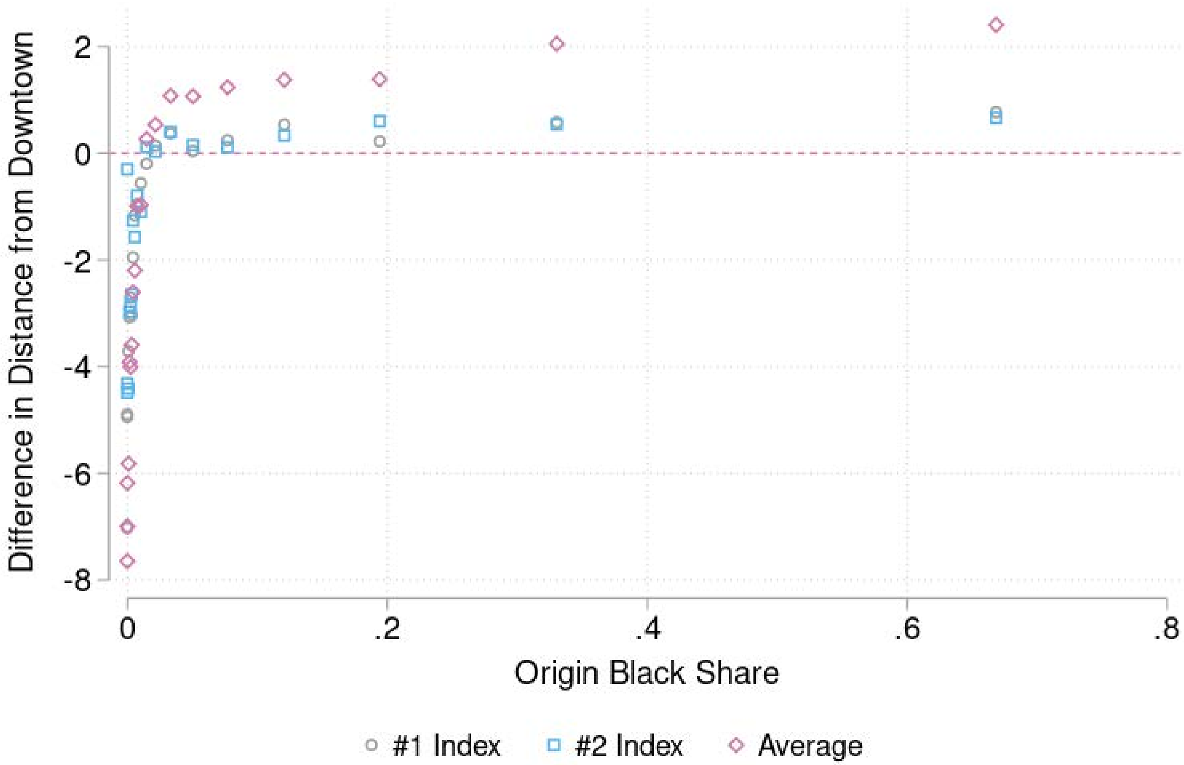
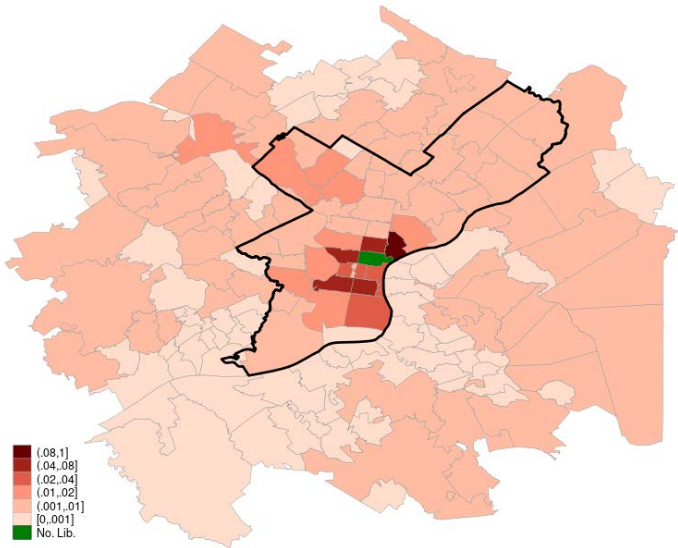
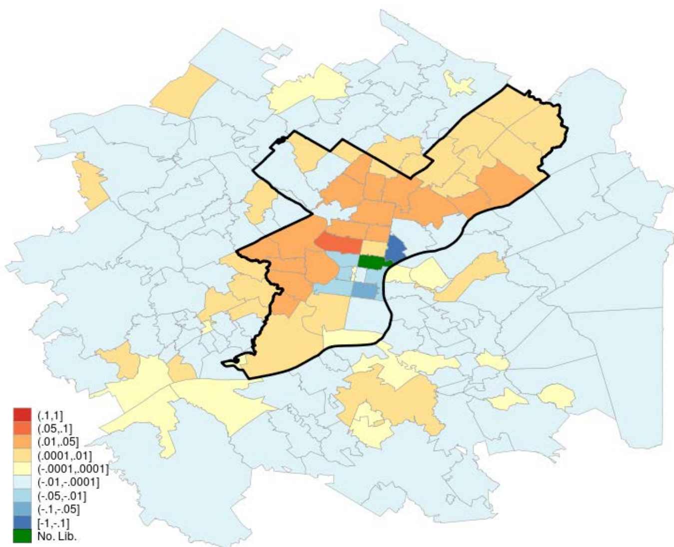


Figure 4: Suburbanization from Black Neighborhoods

Note: Figure reports the average distance from downtown of destination neighborhoods for neighborhoods by origin Black share ventile, as well as the average distance of the highest and second-highest index-value neighborhoods. Positive values of the y-axis signify moves farther from downtown and are reported in miles.



(a) Share of Movers from 19123 (Northern Liberties) by Destination Zip Code



(b) Difference between Share of Black and White Movers from 19123 (Northern Liberties) by Destination Zip Code

Figure 5: Mobility from Northern Liberties

Note: Panel (a) reports the share of movers to a given zip code from 19123 (Northern Liberties). The heat map color buckets are for 0% to 0.1%, 0.1% to 1%, 1% to 2%, 2% to 4%, 4% to 8%, and greater than 8%. Panel (b) reports the difference between the share of Black movers to a given zip code from 19123 minus the share of White movers to a given zip code from 19123. The heat map color buckets are for less than -10 ppts., -10 to -5 ppts., -5 to -1 ppts., -1 to -0.01 ppts., -0.01 ppts. to 0.01 ppts., 0.01 ppts. to 1 ppts., 1 ppt. to 5 ppts., 5 to 10 ppts., and greater than 10 ppts. Thus, red and orange zip codes have greater Black shares of movers, yellow zip codes have nearly identical shares, and blue zip codes have greater White shares. Both panels only show zip codes within 16 miles of City Hall and have the outline of the city of Philadelphia in thick black and Northern Liberties in green. See text for details.

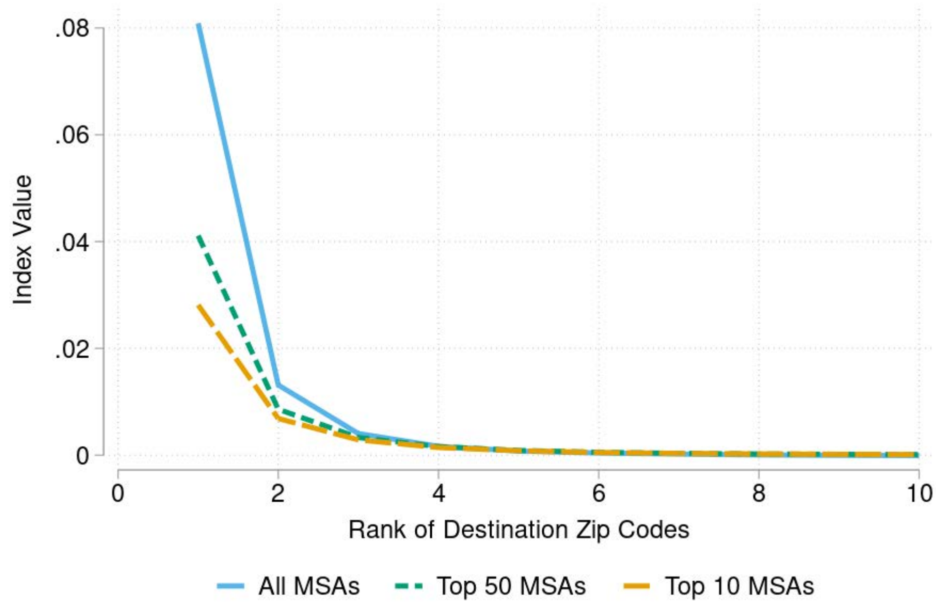
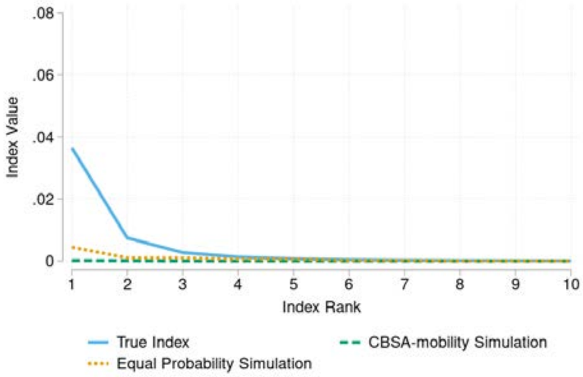
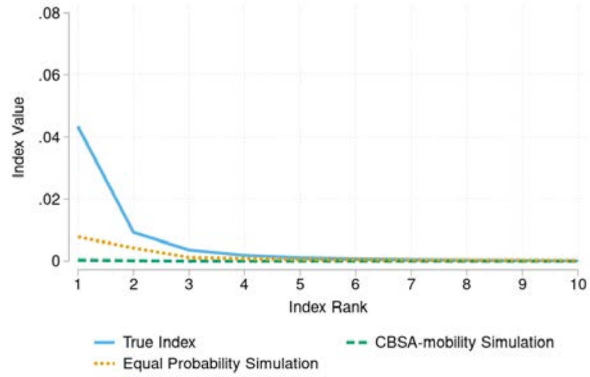


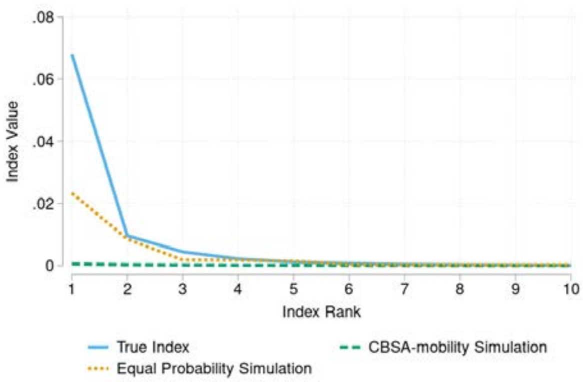
Figure 6: Index Values for Top Neighborhoods
Note: Figure reports the index value for the most connected neighborhoods. See text for details.



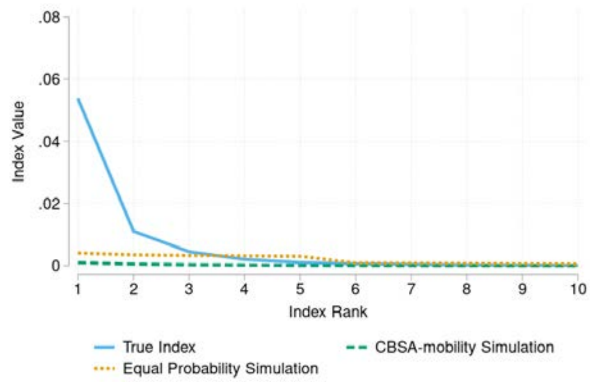
(a) San Francisco-Oakland-Berkeley



(b) Philadelphia-Camden-Wilmington



(c) Cleveland-Elyria



(d) Buffalo-Niagara Falls

Figure 7: Index Rankings by MSA

Note: Figure reports the index value in the reported Metropolitan Statistical Area for ten highest index destination zip codes within the same MSA. The CBSA-mobility simulation and equal probability simulation reports the same data using the 100 data sets simulated using overall mobility probabilities in the MSA and equal mobility probabilities, respectively. See text for details.

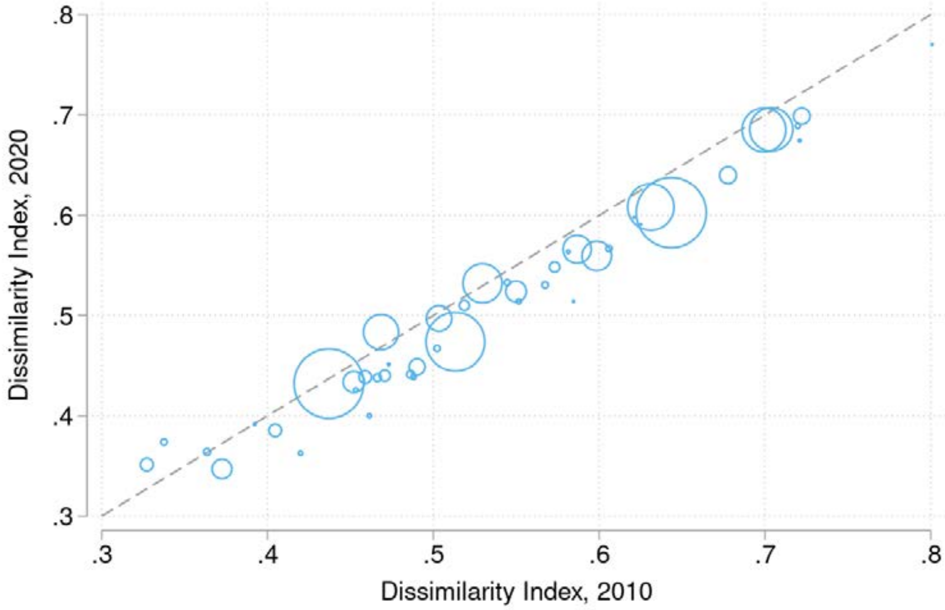


Figure 8: Black-White Dissimilarity Index by MSA, 2010-202 Segregation Impacts
Note: Figure reports the 2010 and 2020 Dissimilarity Index between and Black and White residents of the top 50 MSAs using Census data, showing that intra-MSA segregation is very sticky. Circle sizes reflect the size of the Black population of the MSA. See text for details

6 Tables

Table 1: Summary Statistics

Variable	Mean	Median	Min	Max	SD	N
HPI Pct Change (2010-‘20)	0.445	0.377	-0.348	2.880	0.320	7384
Black Share Change (2010-‘20)	-0.001	0.001	-0.714	0.306	0.037	8576
Hispanic Share Change (2010-‘20)	0.024	0.018	-0.253	0.395	0.033	8576
White Share Change (2010-‘20)	-0.060	-0.058	-0.717	0.414	0.053	8576
Share to Top 1 (2012-‘19)	0.142	0.122	0	1	0.102	8576
Share to Top 2 (2012-‘19)	0.224	0.208	0	1	0.126	8576
Share to Top 5 (2012-‘19)	0.350	0.355	0	1	0.153	8576
Share Out of MSA (2012-‘19)	0.315	0.292	0	1	0.136	8576
Bartik Income Shock	0.185	0.184	0.165	0.213	0.010	8576
Bartik Income Shock	0.170	0.178	0	0.219	0.041	16240*
Share Black (2010)	0.116	0.039	0	0.977	0.187	8576
Share Hispanic (2010)	0.133	0.061	0	1	0.175	8576

Note: Table includes summary statistics for the 8,576 zip codes in our data in the Top 50 MSAs by population. The FHFA HPI, used to calculate HPI Pct Change, does not have complete coverage of these zip codes. Note that the Bartik Income Shock for all 16,240 zip codes in our sample has a standard deviation of 0.041, which is what we use to calculate treatment effects reported in the tables below.

Table 2: Impacts on Origin Neighborhoods

	(1) HPI Pct Change	(2) Black Share Change	(3) Hispanic Share Change	(4) White Share Change
Shock * Share Black	8.331**	-1.629***	0.152	1.519***
	(3.670)	(0.533)	(0.535)	(0.561)
Share Black	-1.278*	0.216**	-0.00885	-0.197*
	(0.679)	(0.0984)	(0.0966)	(0.105)
Shock * Share Hispanic	-2.975	0.163	-0.315	0.184
	(8.516)	(0.409)	(0.817)	(0.864)
Share Hispanic	1.257	-0.0335	0.0330	0.0332
	(1.560)	(0.0759)	(0.155)	(0.164)
Mean of LHS	0.445	-0.001	0.024	-0.060
Treatment Effect (Black)	0.040	-0.008	0.001	0.007
Treatment Effect (Hispanic)	-0.016	0.001	-0.002	0.001
N	7384	8576	8576	8576
R^2	0.722	0.202	0.135	0.181
MSA F.E.	x	x	x	x
Sample	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Dependent variable in Column 1 is the percent change in FHFA zip code level home price index from 2010-2020; Column 2 in the change in the zip code share Black from 2010-2020 using the Census; Column 3 is the change in share Hispanic; Column 4 is the change in share White. "Shock" is a Bartik income shock from 2010 to 2019 using the 2010 and 2019 ACS Samples. "Share Black" and "Share Hispanic" are the share of residents in the zip code who are Black or Hispanic, respectively, using the 2010 Census. "Treatment Effect" refers to the effect of a one standard deviation increase in the Bartik income shock (0.041) for an average share Black zip code (0.113) or share Hispanic zip code (0.133). All specifications include MSA fixed effects. Standard errors are clustered by MSA.

Table 3: Impacts on Origin Neighborhoods - Controlling for Income

	(1) HPI Pct Change	(2) Black Share Change	(3) Hispanic Share Change	(4) White Share Change
Shock * Share Black	7.233**	-1.658**	0.411	1.293**
	(3.265)	(0.702)	(0.400)	(0.567)
Shock*Share Hispanic	-6.341	-0.220	0.218	-0.320
	(9.840)	(0.494)	(0.814)	(0.971)
Shock * ln(Income)	-2.222	-0.126	0.165	-0.157
	(2.343)	(0.259)	(0.224)	(0.331)
Shock * Share 1 Yr Ago	-5.892	-1.262	2.854***	-1.316
	(8.673)	(1.036)	(0.932)	(1.774)
Share Black	-1.171*	0.229*	-0.0582	-0.175
	(0.614)	(0.129)	(0.0742)	(0.107)
Share Hispanic	1.773	0.0427	-0.0643	0.104
	(1.797)	(0.0910)	(0.153)	(0.181)
ln(Income)	0.322	0.0303	-0.0317	0.00680
	(0.442)	(0.0465)	(0.0415)	(0.0612)
Share 1 Yr Ago	0.895	0.221	-0.514***	0.266
	(1.598)	(0.188)	(0.168)	(0.320)
Mean of LHS	0.445	-0.001	0.024	-0.060
Treatment Effect (Black)	0.034	-0.008	0.003	0.006
Treatment Effect (Hispanic)	-0.035	-0.001	0.001	-0.002
N	7302	8403	8403	8403
R^2	0.730	0.209	0.145	0.207
MSA F.E.	x	x	x	x
Sample	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Dependent variable in Column 1 is the percent change in FHFA zip code level home price index from 2010-2020; Column 2 in the change in the zip code share Black from 2010-2020 using the Census; Column 3 is the change in share Hispanic; Column 4 is the change in share White. "Shock" is a Bartik income shock from 2010 to 2019 using the 2010 and 2019 ACS Samples. "Share Black" and "Share Hispanic" are the share of residents in the zip code who are Black or Hispanic, respectively, using the 2010 Census. "ln(Income)" is the natural log of the median income of the zip code using the 2010 ACS. "Share 1 Yr Ago" is share of residents in the zip code who were in a different residence one year ago using the 2010 ACS. "Treatment Effect" refers to the effect of a one standard deviation increase in the Bartik income shock (0.041) for an average share Black zip code (0.113) or share Hispanic zip code (0.133). All specifications include MSA fixed effects. Standard errors are clustered by MSA.

Table 4: Impacts on Movers - Share of Moves to Different Destinations

	(1) Share to Top 1	(2) Share to Top 2	(3) Share to Top 5	(4) Share Out of MSA
Shock * Share Black	2.287***	2.690**	1.663	3.710*
	(0.847)	(1.005)	(1.239)	(1.862)
Share Black	-0.474***	-0.557***	-0.331	-0.867**
	(0.158)	(0.187)	(0.231)	(0.337)
Shock * Share Hispanic	0.576	0.651	0.642	-1.666
	(0.764)	(0.954)	(1.343)	(3.841)
Share Hispanic	-0.132	-0.133	-0.106	0.0891
	(0.144)	(0.178)	(0.254)	(0.717)
Mean of LHS	0.142	0.224	0.350	0.315
Treatment Effect (Black)	0.011	0.013	0.008	0.018
Treatment Effect (Hispanic)	0.003	0.004	0.004	-0.009
N	8576	8576	8576	8576
R^2	0.089	0.100	0.095	0.229
MSA F.E.	x	x	x	x
Sample	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Dependent variable in Column 1 is the share of all moves from 2012-2019 within the MSA but out of the zip code that are to a Top 1 destination as defined by our mobility index; Column 2 is the share to either of the Top 2 destinations; Column 3 to any of the Top 5. Column 4 is the share of all moves out of zip code to a destination out of the MSA. "Shock" is a Bartik income shock from 2010 to 2019 using the 2010 and 2019 ACS Samples. "Share Black" and "Share Hispanic" are the share of residents in the zip code who are Black or Hispanic, respectively, using the 2010 Census. "Treatment Effect" refers to the effect of a one standard deviation increase in the Bartik income shock (0.041) for an average share Black zip code (0.113) or share Hispanic zip code (0.133). All specifications include MSA fixed effects. Standard errors are clustered by MSA.

Table 5: Impacts on Movers - Share of Moves to Different Destinations - Controlling for Income

	(1) Share to Top 1	(2) Share to Top 2	(3) Share to Top 5	(4) Share out of MSA
Shock * Share Black	1.967*	2.157*	-0.0956	2.664
	(0.998)	(1.176)	(1.462)	(2.607)
Shock * Share Hispanic	0.0330	-0.384	-2.073	-3.039
	(1.280)	(1.426)	(1.770)	(4.432)
Shock * ln(Income)	-0.419	-0.957	-2.670*	-1.540
	(0.785)	(1.074)	(1.353)	(1.241)
Shock * Share 1 Yr Ago	-0.772	0.423	3.666	3.462
	(1.840)	(2.614)	(3.810)	(2.916)
Share Black	-0.415**	-0.447**	0.0348	-0.689
	(0.187)	(0.220)	(0.274)	(0.473)
Share Hispanic	-0.0315	0.0712	0.446	0.324
	(0.239)	(0.264)	(0.329)	(0.825)
ln(Income)	0.0759	0.191	0.552**	0.264
	(0.144)	(0.198)	(0.250)	(0.227)
Share 1 Yr Ago	0.164	-0.0990	-0.812	-0.612
	(0.344)	(0.491)	(0.715)	(0.539)
Mean of LHS	0.142	0.224	0.350	0.315
Treatment Effect (Black)	0.009	0.010	-0.0004	0.013
Treatment Effect (Hispanic)	0.0002	-0.002	-0.011	-0.017
N	8403	8403	8403	8403
R^2	0.088	0.102	0.117	0.238
MSA F.E.	x	x	x	x
Sample	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Dependent variable in Column 1 is the share of all moves from 2012-2019 within the MSA but out of the zip code that are to a Top 1 destination as defined by our mobility index; Column 2 is the share to either of the Top 2 destinations; Column 3 to any of the Top 5. Column 4 is the share of all moves out of zip code to a destination out of the MSA. "Shock" is a Bartik income shock from 2010 to 2019 using the 2010 and 2019 ACS Samples. "Share Black" and "Share Hispanic" are the share of residents in the zip code who are Black or Hispanic, respectively, using the 2010 Census. "ln(Income)" is the natural log of the median income of the zip code using the 2010 ACS. "Share 1 Yr Ago" is share of residents in the zip code who were in a different residence one year ago using the 2010 ACS. "Treatment Effect" refers to the effect of a one standard deviation increase in the Bartik income shock (0.041) for an average share Black zip code (0.113) or share Hispanic zip code (0.133).

Table 6: Relative Black Share and Distance between Origin and Destination(s) for All Movers

	(1) Mean Change Black All	(2) Change Black Top 1	(3) Mean Change Dist All	(4) Change Dist Top 1
Shock * Share Black	0.0453	0.285	-162.5*	-99.34
	(2.510)	(2.747)	(95.29)	(63.49)
Share Black	-0.532	-0.311	39.41**	21.60*
	(0.460)	(0.512)	(17.86)	(11.87)
Shock * Share Hispanic	1.813	0.825	58.42	-36.74
	(1.241)	(1.336)	(156.5)	(91.49)
Share Hispanic	-0.281	-0.140	0.0302	9.556
	(0.225)	(0.245)	(29.48)	(17.02)
Mean of LHS	0.003	0.001	-2.52	-0.94
Treatment Effect (Black)	0.0002	0.001	-0.773	-0.472
Treatment Effect (Hispanic)	0.010	0.004	0.319	-0.200
N	8576	7907	8576	7907
R^2	0.803	0.190	0.129	0.046
MSA F.E.	x	x	x	x
Sample	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Dependent variable in Column 1 is the average of the difference between the Black share of the origin zip code and all destination zip codes, weighted by all moves from 2012-2019 to the given destinations. Column 2 is the difference between the Black share of the origin zip code and the Top 1 destination zip code, as defined by our mobility index. Column 3 is the average of the difference between the distance the MSA center of the origin zip code and all destination zip codes, weighted by all moves from 2012-2019 to the given destinations. Column 4 is the difference between the distance to the MSA center of the origin zip code and the Top 1 destination zip code, as defined by our mobility index. Positive numbers indicate the Black share or the distance to MSA center in the destinations is **greater**. “Treatment Effect” refers to the effect of a one standard deviation increase in the Bartik income shock (0.041) for an average share Black zip code (0.113) or share Hispanic zip code (0.133). All specifications include MSA fixed effects. Standard errors are clustered by MSA.

Table 7: Impacts on Black Movers Only

	(1) Share to Top 1	(2) Share to Top 2	(3) Share to Top 5	(4) Share Out of MSA
Shock * Share Black	4.420***	5.162***	4.848***	5.243*
	(1.257)	(1.448)	(1.534)	(2.623)
Share Black	-0.888***	-1.039***	-0.968***	-1.188**
	(0.239)	(0.276)	(0.291)	(0.478)
Shock * Share Hispanic	-0.437	-0.173	-0.647	0.682
	(1.130)	(1.309)	(1.372)	(3.942)
Share Hispanic	0.0414	-0.0137	0.0722	-0.380
	(0.213)	(0.248)	(0.260)	(0.741)
Mean of LHS	0.151	0.240	0.381	0.303
Treatment Effect (Black)	0.021	0.025	0.023	0.025
Treatment Effect (Hispanic)	-0.002	-0.001	-0.004	0.004
N	7795	7795	7795	7958
R^2	0.060	0.071	0.079	0.138
MSA F.E.	x	x	x	x
Sample	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Dependent variable in Column 1 is the share of all moves from 2012-2019 within the MSA but out of the zip code that are to a Top 1 destination as defined by our mobility index; Column 2 is the share to either of the Top 2 destinations; Column 3 to any of the Top 5. Column 4 is the share of all moves out of zip code to a destination out of the MSA. Race is assigned using NamePrism data and the algorithm described above (see text for details). "Shock" is a Bartik income shock from 2010 to 2019 using the 2010 and 2019 ACS Samples. "Share Black" and "Share Hispanic" are the share of residents in the zip code who are Black or Hispanic, respectively, using the 2010 Census. "Treatment Effect" refers to the effect of a one standard deviation increase in the Bartik income shock (0.041) for an average share Black zip code (0.113) or share Hispanic zip code (0.133). All specifications include MSA fixed effects. Standard errors are clustered by MSA.

Table 8: Relative Black Share and Distance between Origin and Destination(s) for Black Movers Only

	(1) Mean Change Black All	(2) Mean Change Distance All
Shock * Share Black	0.354	-133.8*
	(2.360)	(72.63)
Share Black	-0.507	31.44**
	(0.432)	(13.57)
Shock * Share Hispanic	2.528	-33.42
	(2.541)	(118.2)
Share Hispanic	-0.366	14.65
	(0.464)	(22.23)
Mean of LHS	0.023	-1.75
Treatment Effect (Black)	0.002	-0.636
Treatment Effect (Hispanic)	0.014	-0.182
N	7795	7795
R^2	0.567	0.084
MSA F.E.	x	x
Sample	Top 50 MSAs	Top 50 MSAs

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Dependent variable in Column 1 is the average of the difference between the Black share of the origin zip code and all destination zip codes, weighted by moves by Black movers only from 2012-2019 to the given destinations. Column 2 is the average of the difference between the distance the MSA center of the origin zip code and all destination zip codes, weighted by moves from 2012-2019 by Black movers only to the given destinations. Positive numbers indicate the Black share or the distance to MSA center in the destinations is **greater**. Race is assigned using NamePrism data and the algorithm describes above (see text for details). "Treatment Effect" refers to the effect of a one standard deviation increase in the Bartik income shock (0.041) for an average share Black zip code (0.113) or share Hispanic zip code (0.133). All specifications include MSA fixed effects. Standard errors are clustered by MSA.

Table 9: Dissimilarity Index Change and Simulations, 2010-2020

	(1) Black-White	(2) Hispanic-White
Dissimilarity Index, 2010	0.535	0.405
Dissimilarity Index, 2020	0.511	0.387
Simulated Dissimilarity Index, 2020	0.509	0.365
Simulated 2020 Index with 1 SD Income Shock	0.495	0.358
Simulated 2020 Index with top 50 MSA 1 SD Income Shock	0.505	0.363

Note: The first two rows report the average dissimilarity index for the top 50 largest MSAs. The third row uses the estimates in Table 2 to predict the White, Black, and Hispanic shares in every neighborhood, which we force to be between 0% and 100%; we then use the true 2020 zip code populations to estimate the racial/ethnic populations, which we then use to calculate the dissimilarity index. Rows 4 and 5 follow the same methodology but add a one standard deviation income shock to the true income shock for the simulation.

Appendix

A Figures

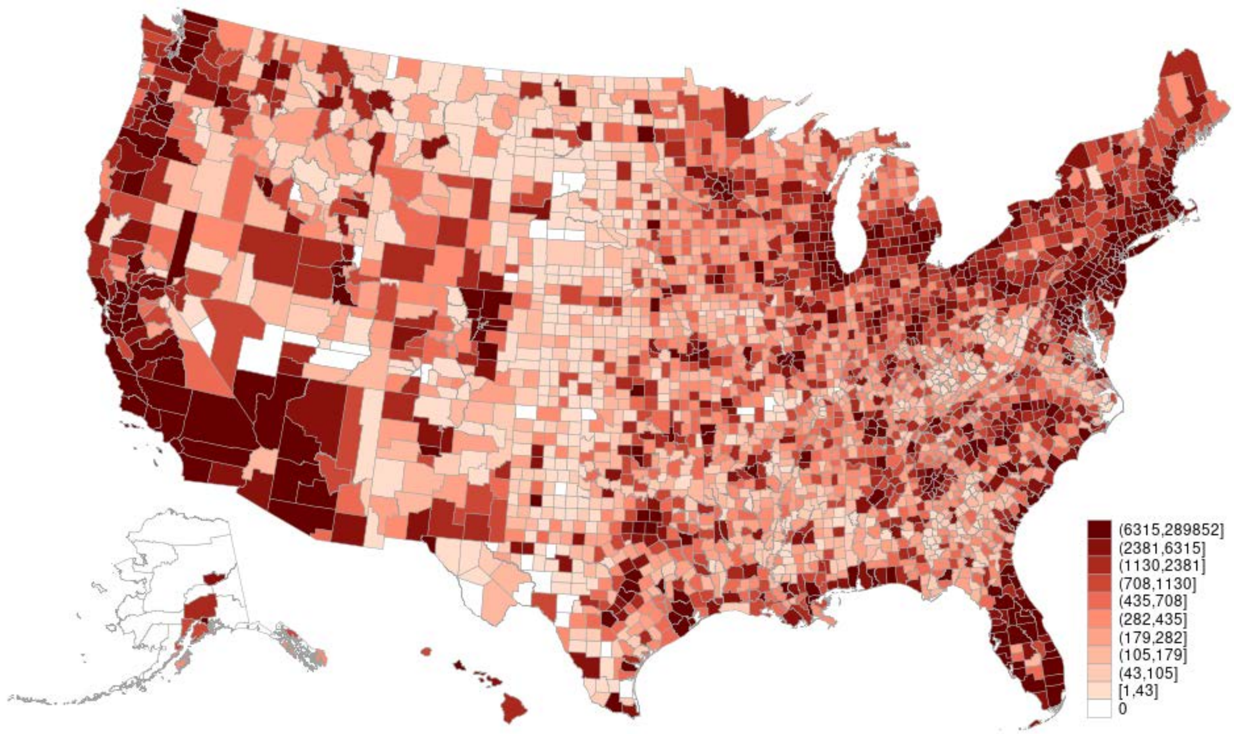
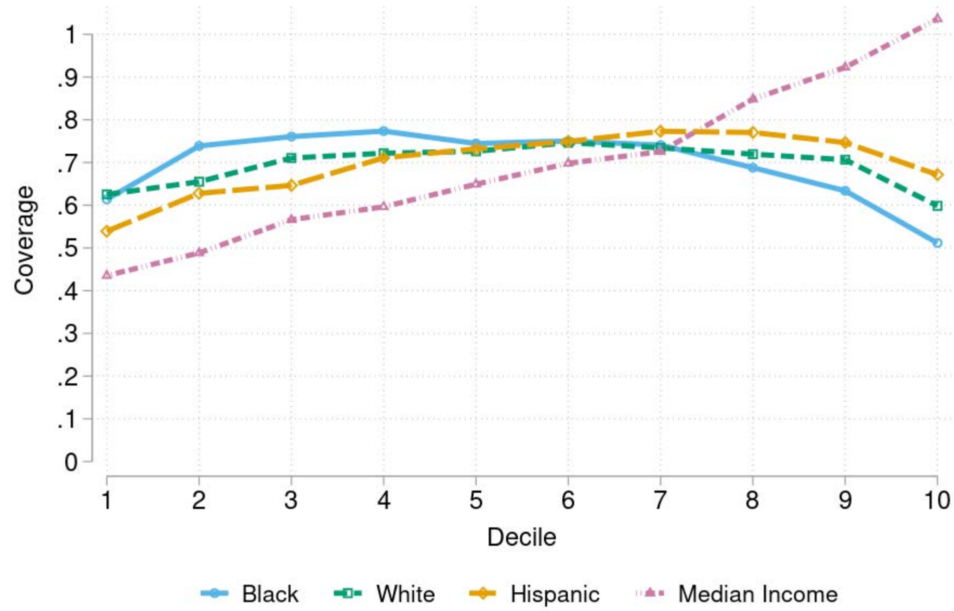
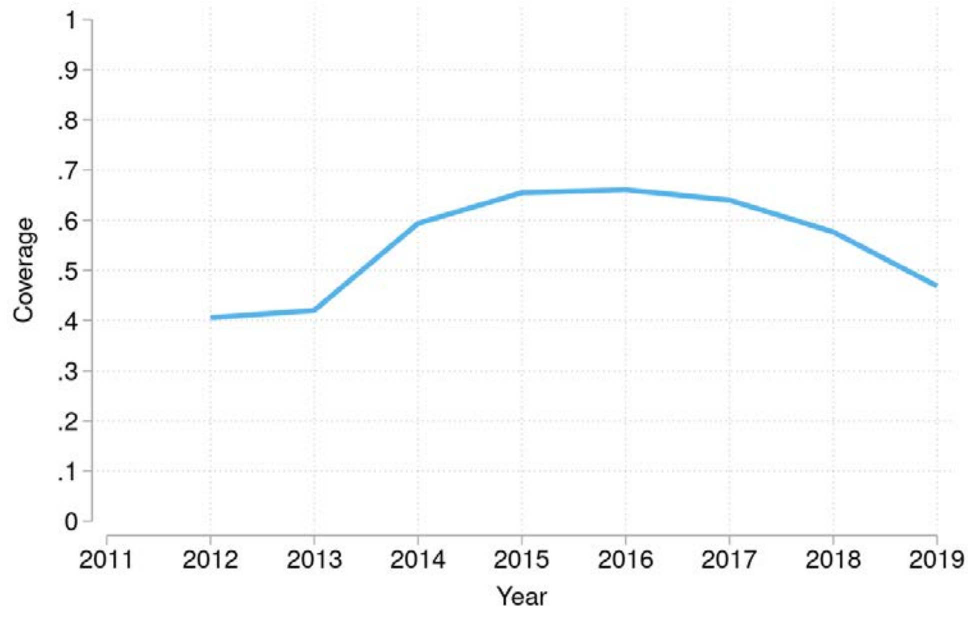


Figure A1: Moves by Origin County, 2015

Note: Map reports the number of moves from counties in the cleaned GrayHair data. Heat map constructed using deciles for counties with at least one move in 2015. See text for details.



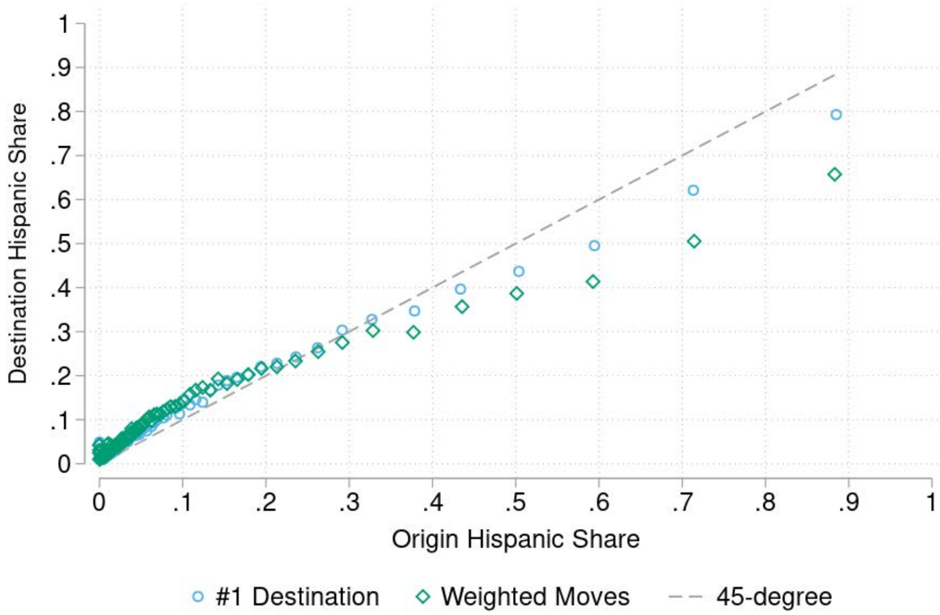
(a) Coverage by characteristic, 2015



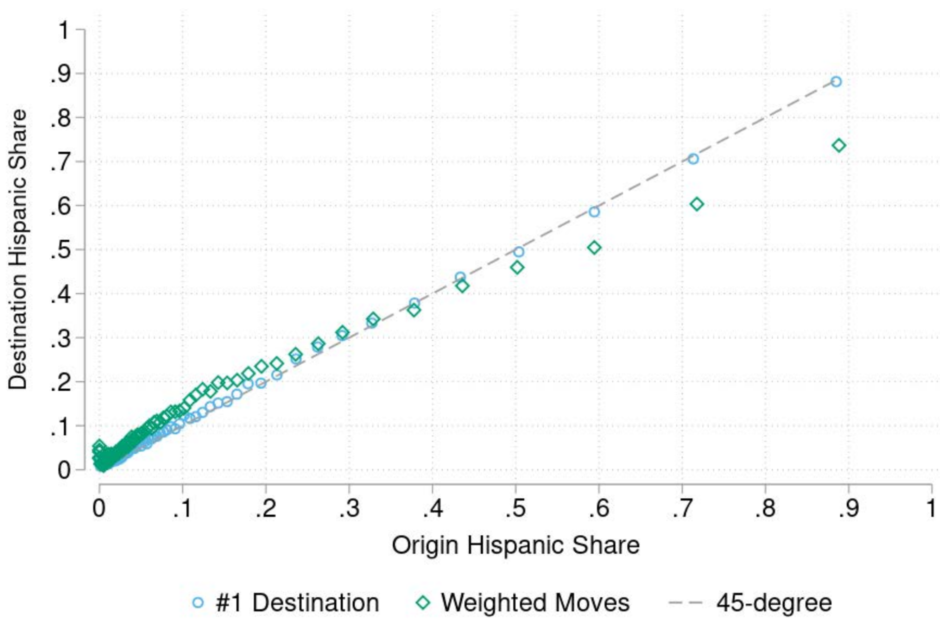
(b) Coverage by year

Figure A2: Coverage of GrayHair Data

Note: Panel (a) benchmarks the number of moves in GrayHair in a Public-Use Microdata Area (PUMA) to the number of people reporting to the American Community Survey that they moved in the past year in a PUMA, and reports the average share for PUMAs by demographic deciles. Average coverage for a given PUMA characteristic is shown by the decile of that same characteristic. Average coverage over time in Panel (b) reports the overall coverage of GrayHair compared to the ACS.



(a) Mobility from Hispanic Neighborhoods, All Movers



(b) Mobility from Hispanic Neighborhoods, Hispanic Movers

Figure A3: Mobility Concentration out of Hispanic Neighborhoods

Note: Panel (a) reports the Hispanic share of the number one and the average Hispanic share of destination neighborhoods by origin neighborhood Hispanic percentile. Panel (b) reports the same but for movers who are Hispanic. The regression estimates of the destination Black share on the origin Black share for #1 destinations and weighted moves are 0.91 and 0.66 in Panel (a) and 0.91 and 0.74 in Panel (b), respectively. See text for details.

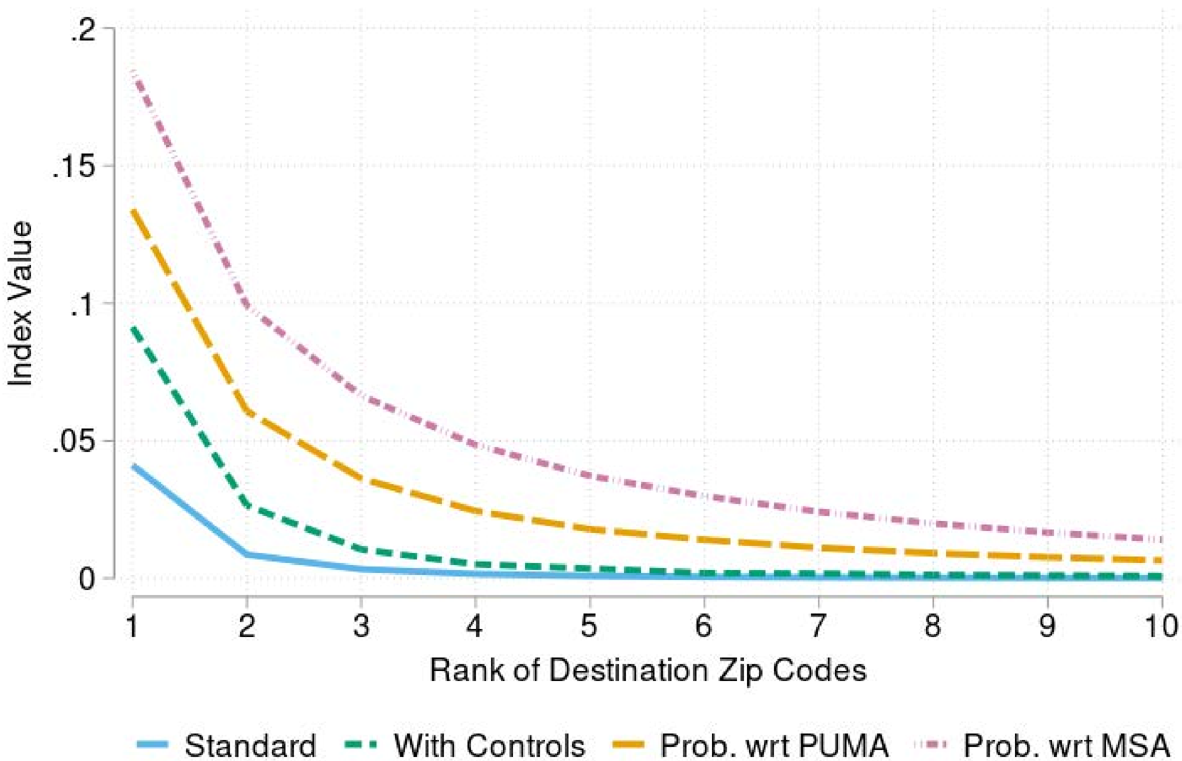


Figure A4: Levels of Alternative Indices

Note: See text for details on the construction of the alternative indices.

B Tables

Table B1: GrayHair Observation Counts

Year	Count
2009	6,216,934
2010	8,139,449
2011	8,412,634
2012	6,065,325
2013	6,294,439
2014	8,889,381
2015	9,826,230
2016	9,881,391
2017	9,573,303
2018	8,763,960
2019	8,442,810

Table B2: Impacts on Origin Neighborhoods, no MSA FEs

	(1) HPI Pct Change	(2) Black Share Change	(3) Hispanic Share Change	(4) White Share Change
Shock * Share Black	11.62	-1.563**	0.569	0.972
	(7.087)	(0.626)	(0.574)	(0.615)
Share Black	-2.048	0.212*	-0.0864	-0.0982
	(1.321)	(0.116)	(0.103)	(0.114)
Shock * Share Hispanic	2.975	0.405	-1.145	1.480
	(16.54)	(0.481)	(0.987)	(1.049)
Share Hispanic	0.483	-0.0840	0.213	-0.237
	(3.043)	(0.0883)	(0.188)	(0.201)
Shock	-5.946**	0.0643	0.618***	-0.958***
	(2.879)	(0.0771)	(0.146)	(0.171)
Mean of LHS	0.445	-0.001	0.024	-0.060
Treatment Effect (Black)	0.055	-0.007	0.003	0.005
Treatment Effect (Hispanic)	0.016	0.002	-0.006	0.008
N	7384	8576	8576	8576
R^2	0.316	0.157	0.040	0.110
MSA F.E.	No	No	No	No
Sample	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Dependent variable in Column 1 is the percent change in FHFA zip code level home price index from 2010-2020; Column 2 in the change in the zip code share Black from 2010-2020 using the Census; Column 3 is the change in share Hispanic; Column 4 is the change in share White. "Shock" is a Bartik income shock from 2010 to 2019 using the 2010 and 2019 ACS Samples. "Share Black" and "Share Hispanic" are the share of residents in the zip code who are Black or Hispanic, respectively, using the 2010 Census. "Treatment Effect" refers to the effect of a one standard deviation increase in the Bartik income shock (0.041) for an average share Black zip code (0.113) or share Hispanic zip code (0.133). These specifications **do not** include MSA fixed effects, and as such the Bartik income shock, which is an MSA-level measure, is included. Standard errors are clustered by MSA.

Table B3: Impacts on Origin Neighborhoods by Demographic Share Quartile

	(1) HPI Pct Change	(2) Black Share Change	(3) Hispanic Share Change	(4) White Share Change
Shock * Black Q2	-1.395	-0.0998	0.0456	0.181
	(1.057)	(0.0754)	(0.143)	(0.248)
Shock * Black Q3	-2.283	-0.444***	-0.135	0.592*
	(1.631)	(0.0980)	(0.193)	(0.306)
Shock * Black Q4	-1.117	-0.750**	-0.238	0.936**
	(1.981)	(0.322)	(0.286)	(0.368)
Black Q2	0.266	0.0175	-0.00412	-0.0435
	(0.191)	(0.0136)	(0.0257)	(0.0446)
Black Q3	0.473	0.0799***	0.0279	-0.119**
	(0.292)	(0.0178)	(0.0344)	(0.0545)
Black Q4	0.330	0.113*	0.0520	-0.155**
	(0.363)	(0.0584)	(0.0511)	(0.0660)
Shock * Hispanic Q2	2.444*	0.223	-0.298*	0.00897
	(1.348)	(0.314)	(0.155)	(0.342)
Shock * Hispanic Q3	4.328**	0.0476	-0.458**	0.411
	(1.887)	(0.272)	(0.203)	(0.473)
Shock * Hispanic Q4	3.444	0.391	-0.346	-0.155
	(2.354)	(0.268)	(0.288)	(0.485)
Hispanic Q2	-0.420*	-0.0331	0.0544*	-0.0161
	(0.238)	(0.0564)	(0.0275)	(0.0605)
Hispanic Q3	-0.760**	0.00676	0.0927**	-0.105
	(0.332)	(0.0482)	(0.0365)	(0.0843)
Hispanic Q4	-0.444	-0.0554	0.0757	0.00455
	(0.434)	(0.0477)	(0.0537)	(0.0872)
Mean of LHS	0.445	-0.001	0.024	-0.060
Treatment Effect (Black)	-0.015	-0.010	-0.003	0.013
Treatment Effect (Hispanic)	0.048	0.005	-0.005	-0.002
N	7384	8576	8576	8576
R^2	0.680	0.114	0.142	0.135
MSA F.E.	No	No	No	No
Sample	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Dependent variable in Column 1 is the percent change in FHFA zip code level home price index from 2010-2020; Column 2 in the change in the zip code share Black from 2010-2020 using the Census; Column 3 is the change in share Hispanic; Column 4 is the change in share White. "Shock" is a Bartik income shock from 2010 to 2019 using the 2010 and 2019 ACS Samples. "Black QX" or "Hispanic QX" is an indicator equal to 1 if the zip code is in the Xth quartile of zip codes for share Black or share Hispanic, respectively, using the 2010 Census. The omitted group is the first quartile (i.e., the lowest share Black/Hispanic zip codes). "Treatment Effect" refers to the effect of a one standard deviation increase in the Bartik income shock (0.041) for an average share Black zip code in the 4th quartile (0.330) or share Hispanic zip code in the 4th quartile (0.339). All specifications include MSA fixed effects. Standard errors are clustered by MSA.

Table B4: Impacts on Movers - Share of Moves to Different Destinations, no MSA FEs

	(1) Share to Top 1	(2) Share to Top 2	(3) Share to Top 5	(4) Share Out of MSA
Shock * Share Black	3.339***	4.130***	3.369***	3.173
	(0.724)	(0.884)	(1.069)	(2.020)
Share Black	-0.650***	-0.795***	-0.615***	-0.739*
	(0.135)	(0.165)	(0.199)	(0.370)
Shock * Share Hispanic	0.165	0.369	0.996	-7.895**
	(1.664)	(2.490)	(2.959)	(3.555)
Share Hispanic	-0.0813	-0.116	-0.202	1.297*
	(0.313)	(0.470)	(0.561)	(0.648)
Shock	-1.669***	-2.331***	-2.884***	1.352
	(0.531)	(0.687)	(0.842)	(1.162)
Mean of LHS	0.142	0.224	0.350	0.315
Treatment Effect (Black)	0.016	0.020	0.016	0.015
Treatment Effect (Hispanic)	0.001	0.002	0.005	-0.043
N	8576	8576	8576	8576
R^2	0.031	0.031	0.024	0.099
MSA F.E.	No	No	No	NO
Sample	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Dependent variable in Column 1 is the share of all moves from 2012-2019 within the MSA but out of the zip code that are to a Top 1 destination as defined by our mobility index; Column 2 is the share to either of the Top 2 destinations; Column 3 to any of the Top 5. Column 4 is the share of all moves out of zip code to a destination out of the MSA. "Shock" is a Bartik income shock from 2010 to 2019 using the 2010 and 2019 ACS Samples. "Share Black" and "Share Hispanic" are the share of residents in the zip code who are Black or Hispanic, respectively, using the 2010 Census. "Treatment Effect" refers to the effect of a one standard deviation increase in the Bartik income shock (0.041) for an average share Black zip code (0.113) or share Hispanic zip code (0.133). These specifications **do not** include MSA fixed effects, and as such the Bartik income shock, which is an MSA-level measure, is included. Standard errors are clustered by MSA. Standard errors are clustered by MSA.

Table B5: Impacts on Movers by Demographic Share Quartile - Share of Moves to Different Destination

	(1) Share to Top 1	(2) Share to Top 2	(3) Share to Top 5	(4) Share Out of MSA
Shock * Black Q2	0.757	0.790	-0.230	1.042
	(0.548)	(0.815)	(1.343)	(1.287)
Shock * Black Q3	1.367**	1.625*	0.631	2.016
	(0.618)	(0.902)	(1.507)	(1.412)
Shock * Black Q4	2.119***	2.559***	1.358	3.341**
	(0.555)	(0.849)	(1.357)	(1.353)
Share Black Q2	-0.145	-0.145	0.0701	-0.224
	(0.0994)	(0.147)	(0.240)	(0.231)
Share Black Q3	-0.269**	-0.311*	-0.0973	-0.420
	(0.113)	(0.164)	(0.270)	(0.253)
Share Black Q4	-0.417***	-0.496***	-0.242	-0.705***
	(0.102)	(0.155)	(0.245)	(0.242)
Shock * Hispanic Q2	0.777	0.772	1.155	-1.622**
	(0.701)	(0.900)	(1.587)	(0.675)
Shock * Hispanic Q3	-0.317	-0.603	0.0953	-2.849**
	(0.614)	(0.849)	(1.441)	(1.214)
Shock * Hispanic Q4	-0.543	-0.891	0.211	-4.061***
	(0.583)	(0.748)	(1.210)	(1.358)
Hispanic Q2	-0.139	-0.132	-0.181	0.301**
	(0.128)	(0.162)	(0.282)	(0.121)
Hispanic Q3	0.0579	0.120	0.0238	0.504**
	(0.112)	(0.155)	(0.257)	(0.215)
Hispanic Q4	0.0984	0.173	-0.00141	0.694***
	(0.107)	(0.138)	(0.217)	(0.244)
Mean of LHS	0.142	0.224	0.350	0.315
Treatment Effect (Black)	0.029	0.035	0.018	0.045
Treatment Effect (Hispanic)	-0.008	-0.012	0.003	-0.056
N	8576	8576	8576	8576
R^2	0.091	0.100	0.104	0.198
MSA F.E.	No	No	No	No
Sample	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs	Top 50 MSAs

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Dependent variable in Column 1 is the share of all moves from 2012-2019 within the MSA but out of the zip code that are to a Top 1 destination as defined by our mobility index; Column 2 is the share to either of the Top 2 destinations; Column 3 to any of the Top 5. Column 4 is the share of all moves out of zip code to a destination out of the MSA. “Shock” is a Bartik income shock from 2010 to 2019 using the 2010 and 2019 ACS Samples. “Black QX” or “Hispanic Xx” is an indicator equal to 1 if the zip code is in the Xth quartile of zip codes for share Black or share Hispanic, respectively, using the 2010 Census. The omitted group is the first quartile (i.e., the lowest share Black/Hispanic zip codes). “Treatment Effect” refers to the effect of a one standard deviation increase in the Bartik income shock (0.041) for an average share Black zip code in the 4th quartile (0.330) or share Hispanic zip code in the 4th quartile (0.339). All specifications include MSA fixed effects. Standard errors are clustered by MSA.

C Simulated Moves Data Set

To create the data set of simulated moves from 2009-2011, we first assume that the number of movers from a given neighborhood is constant; that is we hold fixed the probability of moving out of a neighborhood. We then draw this number of random moves from a distribution with equal probability of moving to a destination zip code. Finally, we drop moves where the origin destination is equal to the destination neighborhood. We refer to this simulated data set of moves as the Equal Probability Simulation data set and use it in Figures 2 and 7 as described in the text.

To create the CBSA-mobility Simulation data set, we first construct an the empirical destination probability among all moves within the MSA. For each origin neighborhood, we do so using a "leave-one-out" methodology, where we consider the probability of moving to any zip code in the MSA excluding moves from the given origin neighborhood. We then draw from this distribution of destination neighborhoods setting the number of draws equal to the number of moves from a given neighborhood, and drop moves where the origin is equal to the destination. As above, we use this data set in Figures 2 and 7.