# Preferences over Race, the Inherent Instability of Neighborhoods in the United States, and the Implications for Public Policy*


Morris A. Davis
*Rutgers University*
mdavis@business.rutgers.edu

Jesse Gregory
*University of Wisconsin - Madison*
jmgregory@ssc.wisc.edu

Daniel A. Hartley
*Federal Reserve Bank of Chicago*
Daniel.A.Hartley@chi.frb.org



November 10, 2022



**Abstract**

We estimate the parameters of a dynamic location-choice model in which households care about exogenous amenities of neighborhoods and the endogenous racial composition of residents. We use a new Bartik-style isntrument to estimate preferences for the racial composition of neighborhoods and find that many households have preferences characterized by homophily: For most black, hispanic and white households in our data, the utility of their current neighborhood will increase if the share of households of their race in the neighborhood increases. Preferences over race are sufficiently strong such that even when the model is engineered to be in a steady state consistent with the current data, a perturbation in the expected demographic mix in one neighborhood in a metro area can cause widespread and large changes to many neighborhoods. This result, corroborated by model simulations, suggest that seemingly minor public policies that shift the expected demographic composition of neighborhoods have the potential to cause a significant re-sorting of the population.



*JEL* Classification Numbers: Insert classifications here
*Keywords*: Race, Neighborhood Composition, Public Policy


*This paper was previously circulated with the title "The Long-Run Effects of Low-Income Housing on Neighborhood Composition." The views expressed herein are those of the authors and do not necessarily represent those of the Federal Reserve Bank of Chicago or the Federal Reserve System.

# 1 Introduction

The racial composition of neighborhoods can change quickly. A well known example involves the Weequahic neighborhood of Newark, NJ, birthplace of celebrated novelist Philip Roth. In 1960, 5 Census tracts comprising most of the population of the Weequahic neighborhood were 99% white and 1% black; by 1970, those same tracts were 18% white and 82% black.[1] A more recent example is the Columbia Heights neighborhood of Washington, DC.[2] In 2000, Columbia Heights was 6% white, 58% black and 33% hispanic. By 2020, these percentages had changed to 39% white, 28% black and 25% hispanic.

We try to understand the extent to which preferences over the racial composition of neighborhoods influences where people choose to live, and, by extension, how households may react to policies that change the expected racial composition of neighborhoods. To do this, we estimate a dynamic, forward-looking model of neighborhood choice. We allow households to have preferences over exogenous intrinsic features of neighborhoods as well as preferences over the endogenous demographic composition of neighborhoods. We estimate the parameters of this model using a large and nationally representative data set for neighborhoods located in 141 metro areas in the United States.

Identification of household preferences over the demographic composition of neighborhoods using location choice data is confounded by the presence of location-specific amenities that may be unobserved, and the valuation of these amenities that may differ by race. For example, suppose black households and white households tend to live in different neighborhoods; and, the neighborhoods that are predominantly black have mostly pine trees and the neighborhoods that are predominantly white have mostly spruce trees. Neighborhoods may be segregated either because black households prefer pine trees and white households prefer spruce trees; because black households prefer to have black neighbors and white households prefer to have white neighbors; or both.

Therefore, to identify preferences over demographic composition of neighborhoods we need to find an instrument that is correlated with racial shares at the neighborhood level but uncorrelated with local amenities such as the prevalence of pine or spruce trees. We use an instrumental-variables approach in the style of Bartik (1991). We believe we are the first to use a Bartik instrument to estimate preferences various types of households have over white, black, and hispanic racial shares of neighborhoods.

[1] We derive these statistics from data reported in the 1960 and 1970 Census of Population and Housing, published by the U.S. Department of Commerce. We only consider tracts 44, 45, 46, 47 and 49; typically the definition of the Weequahic neighborhood also includes tracts 48.01 and 51. For reference, the population of the 4 included tracts was 20,096 in 1960 and 26,598 in 1970; when tracts 48.01 and 51 are included, the population rises to 27,954 in 1960 and 34,041 in 1970.

[2] These are Census tracts 28.01, 28.02, 29, 30, 31, 35, 36 and 37 (2000) or 37.01 and 37.02 (2020).

The instrument exploits the fact that there is variation, across metropolitan areas, in the metro-area shares of black, hispanic, and white households; therefore, metro areas with relatively high shares of black households overall are predicted to have relatively high shares of black households in neighborhood in these metros.

Our IV procedure works as follows. First, we estimate a multinomial logit that predicts the probability any given type of household[3] lives in a given neighborhood (Census tract throughout) based only on that neighborhoods's within-metro income percentile ranking. This logit pools all data from all metro areas. For any given type of household the predicted mapping of location choices to the income percentile ranking will not vary across metro areas.

Next, we predict the distribution of where all households will live within each metro area. This enables us to compute predicted racial shares of black, hispanic and white households in each neighborhood in each metro area. Conditional on the income percentile of the neighborhood, variation in predicted racial shares is entirely attributable to variation in the distribution of types of households across metro areas. Thus, conditional on income percentile, predicted racial composition of each neighborhood should be orthogonal to local amenities (i.e. the prevalence of pine versus spruce trees), implying the instrument is valid.

Before continuing, we should be clear about what we can and cannot identify. We cannot identify pure preferences over race if certain types of households "bring along" certain types of amenities. In this case, we would estimate preferences for the bundle of race and the amenities that are brought with race. For example, if pine and spruce trees happen to be permanently attached to certain neighborhoods, then our procedure should uncover preferences for the racial composition of neighborhoods independent of the prevalence of pine and spruce trees. On the other hand, if black households plant pine trees wherever they live, and white households plant spruce trees wherever they live, then our procedure estimates relative preferences of the bundle (black households, pine trees) and (white households, spruce trees) rather than preferences for race directly.

Obviously, we are tip-toeing around the fact that local governments may systematically underfund policing or schools in non-white neighborhoods. If school quality and public safety are policy choices that depend on the racial composition of neighborhoods, and not intrinsic function of neighborhoods in which non-white residents tend to live, our estimates of racial preferences will also contain preferences for amenities that follow racial shares around due to decisions of local governments. In some contexts understanding preferences for race separately from preferences for amenities that follow race (for whatever reason) may be

[3] We separate households in our data into mutually exclusive and exhaustive types by race, age (young, middle aged, and old), homeonwership status (rent or own), and credit score terciles.

important, but for our analysis we do not think this distinction matters.

Perhaps not surprising, we estimate that many – but not all – households have preferences exhibiting homophily, that is for many households utility in their chosen neighborhood increases if the share of same-race households in the neighborhood increases. To give a sense of the size of these preferences, for the average black household in our data, we find that if the share of black households in their neighborhood increases by 1 percentage point, utility increases by approximately the same as if rental prices decline by 3 pecent. For the average white household in our data, if the share of black households increases by 1 percentage point utility declines by about the same amount as if rental prices increase by approximately 1 percent.

Before continuing, we wish to take a step back. We started this research project with the intention of using a location-choice model to predict how households may endogenously resort in response to a place-based policy such as the Moving to Opportunity experiment that may alter expectations on the racial mix of neighborhoods (Davis, Gregory, Hartley, and Tan, 2021). After estimating model parameters, we noticed in initial counterfactual policy simulations that small changes to the expected demographic composition of neighborhoods frequently yielded a very large resorting of the population across many neighborhoods. That led us to think that before we use the model to predict the results of any counterfactual policy, we need to understand if the model is "stable" (in a way we define next) at the current data.

Of course, the current data almost certainly do not reflect a stable steady state as the unconditional demographic mix of the United States is changing. For example, in 2020 hispanic persons accounted for 18.7% of the U.S. population in 2020, up from 6.5% in 1980. Additionally, the location choices of households may not be consistent with the existing demographic composition of neighborhoods, as pointed out by Caetano and Maheshri (2021). For example, hispanic households may be moving into some majority-black neighborhoods at a rate that implies the hispanic share of households is increasing and black share of households is decreasing relative to current data.

With these important caveats in mind, we wish to give the model its best shot to see if, absent long-run demographic change, the model predicts the current demographic composition of neighborhoods in the United States is stable. To do this, and different from the exercise in Caetano and Maheshri (2021), for each type of household and in each neighborhood in our data, we allow for exogenous "births and deaths" in every period such that the model-predicted distribution of location choices for each type of household in each neighborhood, once we add in these births and deaths, matches the distribution of types across neighborhoods in the current data.[4] With births and deaths included, the model has the feature that the expected future demographic composition of each neighborhood is always exactly equal to the current demographic composition.

Given this definition of a steady state that is consistent with the current data, we next need to take a stand on how households update expectations when expectations are moved away from the steady state. Simplifying a bit, here is our procedure. First, we feed into the model a set of household expectations about racial composition in every tract. Then, we generate household location decisions given those assumptions. Finally, we set new expectations equal to the new racial composition in every tract implied by those decisions.

With that in mind, we evaluate the stability of the racial composition of neighborhoods in the data by computing the eigenvalues of the model at the steady state implied by the data, as we have defined it. We do this for every MSA. If all the eigenvalues in an MSA are less than 1, then expectations of racial composition return back to the data after a small perturbation to those expectations. If at least one eigenvalue is greater than 1, expectations of the racial composition at least one neighborhood does not revert back to the data after a small perturbation.

Our findings are stark. Only one metro area out of 141 has all its eigenvalues less than 1. Many neighborhoods in most MSAs are unstable: For the median MSA, more than 45% of the eigenvalues are larger than 1. We show the fundamental instability of the model arises because households have very strong preferences over the racial composition of their neighborhood. The MSA-wide black share and MSA-wide hispanic share of the population strongly predict the percentage of eigenvalues in the MSA that are larger than 1. Additionally, when we shrink preferences over race by multiplying all coefficients on race in utility by 0.25 and then recompute eigenvalues, the number of eigenvalues larger than 1 drops dramatically: At the median MSA only 3% of the eigenvalues are larger than 1.[5]

In summary, preferences over demographic composition of neighborhoods are so strong that, even without long-run demographic change, and even assuming exogenous "births and deaths" to keep the demographic composition of neighborhoods constant at the current data, a slight perturbation to expected demographic composition in one neighborhood can basically blow up the steady state. Therefore, it should not be surprising that we document that this fundamental instability suggests very small public policies can be destabilizing if these policies change the expected racial composition of neighborhoods.

[4]Note that our model has households aging and dying; dditionally, the model does not include across-MSA moves. In a very reduced-form sense, births and deaths account for these features and does so in such a way as to mechanically ensure the data are in a steady state.

[5]In this exercise, we adjust exogenous amenities such that the overall utility of each neighborhood does not change at the current data.

Specifically, we consider a small policy that affects many neighborhoods: Metro area by metro area, in all tracts that have had at least some low-income housing built since 2000, we add a one-time, unexpected, additional 10% low-income housing units in those tracts (assuming the new residents have a different demographic mix than existing residents) and then simulate the new steady state of each metro area. We measure two things in these simulations relative to the baseline steady state implied by our data. First, the number of tracts in a metro area where either the black share of households or hispanic share of households changes by 5 percentage points or more, and second, changes to metro-area black-white and hispanic-white segregation indexes.

The punchline from model simulations is that neighborhoods change a lot and metro areas tend to become much more segregated as a result of this experiment. At the median MSA, 60% of tracts experience a change in black- or hispanic- share of at least 5 percentage points in the new steady state relative to the old steady state. Further, the black-white and hispanic-white segregation indexes jump, such that the median metro area become much more segregated.

These results are all driven by the very strong preferences for race that we estimate. When we repeat the exercise after multiplying all coefficients on race in utility by 0.25, changes in demographic composition are quite small. At the median MSA, in the new steady state less than 6 percent of tracts experience a change in black- or hispanic- share of greater than 5 percentage points, the hispanic-white segregation index does not change, and the black-white segregation index declines a little.

Practically speaking, our results imply that households preferences for the race of their neighbors are sufficiently strong that well intentioned policies that change the racial composition of neighborhoods may ultimately destabilize neighborhoods and increase segregation. Of course, these changes may take a long time but recent history also tells us that neighborhood composition can change quickly.

# 2 Household Decision Model

We model the system of demand for neighborhoods by considering the decision problem of a household head deciding where his or her family ("household") should live. As in Kennan and Walker (2011) Bayer, McMillan, Murphy, and Timmins (2015), and Davis, Gregory, Hartley, and Tan (2021) we model location choices in a dynamic discrete choice setting. Each year, the household, which is of type $\tau$, can choose to live in one of $J$ locations. Denote $j$ as the household's current location. We write the value to the household, $V_t^{\tau}(\ell \mid j)$, of choosing to live in location $\ell$ in year $t$ given a current location of $j$ and current value of a shock $\epsilon_\ell$ (to be explained later) as

$$V_t^\tau(\ell \mid j, \epsilon_\ell) \;=\; u_t^\tau(\ell \mid j, \epsilon_\ell) + \beta \sum_{\tau'} \gamma^{\tau,\tau'} E_t\left[V_{t+1}^{\tau'}(\ell)\right]$$

In the above equation $u_t^\tau(\ell \mid j, \epsilon_\ell)$ is the flow utility in year $t$ to the household of choosing to live in location $\ell$ given a current location of $j$ and current value of a shock $\epsilon_\ell$; $\beta$ is the discount factor on future expected utility; $\gamma^{\tau,\tau'}$ is the probability that the household becomes type $\tau'$ next year given they are type $\tau$ this year; and $E_t\left[V_{t+1}^{\tau'}(\ell)\right]$ is the expected value in year $t+1$ of a type $\tau'$ household of having chosen to live in neighborhood $\ell$ today. The $t$ subscripts explicitly allow that flow utility and expectations may change over time.

We assume $u_t^\tau(\ell \mid j, \epsilon_\ell)$ is as follows

$$u^\tau(\ell \mid j, \epsilon_\ell) \;=\; \delta_{\ell,t}^\tau \;-\; \kappa^\tau \cdot 1_{\ell \neq j} \;+\; \epsilon_\ell$$

$\delta_{\ell,t}^\tau$ is the flow utility a type $\tau$ household receives in year $t$ from living in neighborhood $\ell$, inclusive of tastes for rents, neighborhood demographics, and any amenities or natural advantages the neighborhood provides; $\kappa^\tau$ are the fixed costs (utility and financial) a household of type $\tau$ must pay when it moves to a different neighborhood i.e. when $\ell \neq j$; $1_{\ell \neq j}$ is an indicator function that is equal to 1 if location $\ell \neq j$ and 0 otherwise; and $\epsilon_\ell$ is a random shock that is known at the time of the location choice. $\epsilon_\ell$ is assumed to be iid across locations, time and people. $\epsilon_\ell$ induces otherwise identical households living at the same location at the same time to optimally choose different future locations.

Denote $\epsilon_1$ as the shock associated with location 1, $\epsilon_2$ as the shock with location 2, and so on. In each period after the vector of $\epsilon$ are revealed (one for each location), households choose the location that yields the maximal value

$$V_t^\tau(j \mid \epsilon_1, \epsilon_2, \ldots, \epsilon_J) \;=\; \max_{\ell \in 1,\ldots,J} V_t^\tau(\ell \mid j, \epsilon_\ell) \tag{1}$$

$E_t\left[V_{t+1}^{\tau'}(j)\right]$ is the expected value of (1) for households of type $\tau'$, where the expectation is taken at year $t$ with respect to the vector of $\epsilon$ and the vector of $\delta_{j',t+1}^{\tau'}$ for $j' \in 1, \ldots, J$ and for all types $\tau'$, in the event these values change over time.

Holding $\delta_{j,t}^\tau$ fixed for all $j$ and $\tau$ for expositional convenience (such that the expected values are time-invariant and time subscripts can be removed) when the $\epsilon$ are assumed to be drawn i.i.d. from the Type 1 Extreme Value Distribution, the expected value function

$E[V^{\tau}(j)]$ has the functional form

$$E[V^{\tau}(j)] \;=\; \log\left\{\sum_{\ell=1}^{J}\exp\widetilde{V}^{\tau}(\ell \mid j)\right\}+\zeta \tag{2}$$

where $\zeta$ is equal to Euler's constant and

$$\widetilde{V}^{\tau}(\ell \mid j) \;=\; \delta_{\ell}^{\tau} \;-\; \kappa^{\tau}\cdot 1_{\ell\neq j} \;+\; \beta\sum_{\tau'}\gamma^{\tau,\tau'}E\left[V^{\tau'}(\ell)\right] \tag{3}$$

That is, the tilde symbol signifies that the shock $\epsilon_{\ell}$ has been omitted.

We use the approach of Hotz and Miller (1993) and employed by Bishop (2012) to generate a likelihood function. This approach does not require that we solve for the value functions. Instead, it can be shown that the log probabilities that choices are observed are simple functions of model parameters $\delta_j^{\tau}$, $\kappa^{\tau}$, $\beta$ and of observed choice probabilities. In other words, a likelihood over choice probabilities observed in data can be generated without solving for value functions.

To see this, start by noting the log of the probability that location $\ell$ is chosen by type $\tau$ given a current location of $j$, call it $p^{\tau}(\ell \mid j)$, has the solution

$$p^{\tau}(\ell \mid j) \;=\; \widetilde{V}^{\tau}(\ell \mid j) \;-\; \log\left\{\sum_{\ell'=1}^{J}\exp\left[\widetilde{V}^{\tau}(\ell' \mid j)\right]\right\} \tag{4}$$

Subtract and add $\widetilde{V}^{\tau}(k \mid j)$ to the right-hand side of the above to derive

$$p^{\tau}(\ell \mid j) \;=\; \widetilde{V}^{\tau}(\ell \mid j)-\widetilde{V}^{\tau}(k \mid j) \;-\; \log\left\{\sum_{\ell'=1}^{J}\exp\left[\widetilde{V}^{\tau}(\ell' \mid j)-\widetilde{V}^{\tau}(k \mid j)\right]\right\} \tag{5}$$

Note that equation (3) implies

$$\begin{aligned}&\widetilde{V}^{\tau}(\ell \mid j)-\widetilde{V}^{\tau}(k \mid j)\\ &=\delta_{\ell}^{\tau}-\delta_{k}^{\tau} \;-\; \kappa^{\tau}\left[1_{\ell\neq j}-1_{k\neq j}\right] \;+\; \beta\sum_{\tau'}\gamma^{\tau,\tau'}\left\{E\left[V^{\tau'}(\ell)\right]-E\left[V^{\tau'}(k)\right]\right\}\end{aligned} \tag{6}$$

But from equation (2),

$$E\left[V^{\tau'}(\ell)\right]-E\left[V^{\tau'}(k)\right] \;=\; \log\left\{\sum_{\ell'=1}^{J}\exp\widetilde{V}^{\tau'}(\ell' \mid \ell)\right\}-\log\left\{\sum_{\ell'=1}^{J}\exp\widetilde{V}^{\tau'}(\ell' \mid k)\right\}$$

Now note that equation (4) implies

$$
\begin{aligned}
p^{\tau'}(k \mid \ell) &= \widetilde{V}^{\tau'}(k \mid \ell) - \log\left\{\sum_{\ell'=1}^{J}\exp\left[\widetilde{V}^{\tau'}(\ell' \mid \ell)\right]\right\} \\
p^{\tau'}(k \mid k) &= \widetilde{V}^{\tau'}(k \mid k) - \log\left\{\sum_{\ell'=1}^{J}\exp\left[\widetilde{V}^{\tau'}(\ell' \mid k)\right]\right\}
\end{aligned}
$$

and thus

$$
\log\left\{\sum_{\ell'=1}^{J}\exp\left[\widetilde{V}^{\tau'}(\ell' \mid \ell)\right]\right\} - \log\left\{\sum_{\ell'=1}^{J}\exp\left[\widetilde{V}^{\tau'}(\ell' \mid k)\right]\right\}
$$

is equal to

$$
\begin{aligned}
& \widetilde{V}^{\tau'}(k \mid \ell) - \widetilde{V}^{\tau'}(k \mid k) - \left[p^{\tau'}(k \mid \ell) - p^{\tau'}(k \mid k)\right] \\
= & -\kappa^{\tau'} \cdot 1_{\ell \neq k} - \left[p^{\tau'}(k \mid \ell) - p^{\tau'}(k \mid k)\right]
\end{aligned}
$$

The last line is quickly derived from equation (3). Therefore,

$$
E\left[V^{\tau'}(\ell)\right] - E\left[V^{\tau'}(k)\right] = -\left[p^{\tau'}(k \mid \ell) - p^{\tau'}(k \mid k) + \kappa^{\tau'} \cdot 1_{\ell \neq k}\right]
$$

and equation (6) has the expression

$$
\begin{aligned}
& \widetilde{V}^{\tau}(\ell \mid j) - \widetilde{V}^{\tau}(k \mid j) \\
& = \delta_{\ell}^{\tau} - \delta_{k}^{\tau} - \kappa^{\tau}\left[1_{\ell \neq j} - 1_{k \neq j}\right] - \beta\sum_{\tau'}\gamma^{\tau,\tau'}\left[p^{\tau'}(k \mid \ell) - p^{\tau'}(k \mid k) + \kappa^{\tau'} \cdot 1_{\ell \neq k}\right]
\end{aligned} \tag{7}
$$

In estimation, we assume that value functions and expectations are fixed in our sample period.[6] We can therefore use (7) directly in equation (5) for use in the likelihood function, treating observed choice probabilities for $p^{\tau'}(k \mid \ell)$ and $p^{\tau'}(k \mid k)$ for all $\ell$, $k$ and $\tau'$ as data.

# 3 Data and Likelihood

Like Davis, Gregory, Hartley, and Tan (2021), we estimate the model using panel data from the FRBNY Consumer Credit Panel / Equifax. The panel is comprised of a 5% random sample of U.S. adults with a social security number, conditional on having an active credit file, and any individuals residing in the same household as an individual from that initial

[6] Due to data limitations, we combine data across multiple years when estimating probabilities and preference parameters. For this reason, we assume value functions are fixed.

5% sample.[7] For years 1999 to **XXX**, the database provides a quarterly record of variables related to debt: Mortgage and consumer loan balances, payments and delinquencies and some other variables we discuss later. The data does not contain information on basic demographics like race, education, or number of children and it does not contain information on income or assets although it does include the Equifax Risk Score$^{TM}$ which provides some information on the financial wherewithal of the household as demonstrated in Board of Governors of the Federal Reserve System (2007). Most important for our application, the panel data includes in each period the current Census block of residence. To match the annual frequency of our location choice model, we use location data from the first quarter of each calendar year. In each year, we only include people living in in MSAs – if, for example, a household moves from an eligible MSA to a rural area, that household-year observation is not included in the estimation sample. Additionally, if a household moves to a different MSA, the household-year observation of the move is not included in the estimation sample (but the years before and after the across-MSA move are included). The panel is not balanced, as some individuals' credit records first become active after 1999. The total number of person-year observations in the sample is **XXX**.

We sort households into 54 mutually exclusive types: by age of the head of the household (young, middle, old), by housing tenure status (renter, owner), by credit score (low, middle, high), and by race (black, hispanic, white/other). Except for race, a household's type can stochastically change over time. Borrowing a trick from overlapping generations models in macroeconomics to conserve on state variables, we assume households assume they age up (i.e. low to middle, middle to high) or die (high to death) with a 5% probability each year. Conditional on age and race, we estimate the annual 6x6 matrix of transition probabilities of housing tenure status and credit score using the Equifax data pertaining to our estimation sample.

From the Equifax data, we classify a household as young if the age of the household head is between 25-44, middle aged if 45-64, and old if 65 and older. We classify the household as a homeowner if the household has a mortgage and a renter if not. Finally, we classify a household as having a low credit score if the Equifax Risk Score$^{TM}$ of the household head is less than or equal to 599, middle credit score if between 600 and 720 inclusive, and high credit score if greater than or equal to 721.

We do not observe race in the Equifax data. However we know the racial distribution in the Census block in which the household is first observed, and we use this distribution to

[7] The data include all individuals with 5 out of the 100 possible terminal 2-digit social security number (SSN) combinations. While the leading SSN digits are based on the birth year/location, the terminal SSN digits are essentially randomly assigned. A SSN is required to be included in the data and we do not capture the experiences of illegal immigrants.

integrate out uncertainty over the household's race. Let the subscript $r$ denote a household's race, with $r = 1$ a black household, $r = 2$ a hispanic household and $r = 3$ a white/other household which we call a "white" household from this point forward. Let $Pr_i(r)$ denote the probability that household $i$ is of race $r$ given the first observed Census block of residence of that household, i.e. $Pr_i(1)$ is the percentage of residents in that block that are black, $Pr_i(2)$ is the percentage of residents in that block that are hispanic, and $Pr_i(3)$ is the percentage of residents in that block that are white. Then, the likelihood for a given household $i$ that is observed from periods $t = 1, \ldots, T$ can be written as

$$\mathcal{L}_i \;=\; \sum_{r=1}^{3} Pr_i(r) \prod_{t=1}^{T} p^{\tau_t(r)}(\ell_t \mid j_t)$$

where $j_t$ is the household's starting location at year $t$, $\ell_t$ is the household's choice of new location, and $\tau_t(r)$ is the household's type at year $t$ given the household is assumed to be of race $r$. We find the parameters of the model that maximize the log likelihood of the sample, $\sum_i \log \mathcal{L}_i$.

# 4 Parameter Estimates and Identification

## 4.1 Specification of Flow Utility

So far, the key inputs to the model that we have discussed are type-specific moving costs $\kappa^\tau$ and the type specific flow utility of living in any neighborhood $\delta_j^\tau$ for all $j = 1, \ldots, J$. In estimation, we assume $\delta_j^\tau$ is fixed. In the counterfactual simulations, we have in mind that for each type and neighborhood, $\delta_j^\tau$ has a fixed component that we will call "amenities" and a component that may change over time if the racial mix of the neighborhood changes or if the price of housing changes. We specify

$$\delta_j^\tau \;=\; -a_r^\tau \log r_j \;+\; f^\tau\left(S_j^b, S_j^h\right) \;+\; A_j^\tau \tag{8}$$

In the above equation $A_j^\tau$ are the fixed type-specific amenities associated with neighborhood $j$, $r_j$ is the price per unit of housing in neighborhood $j$, and $a_r^\tau$ governs how utility changes with respect to changes in house prices. $f^\tau\left(S_j^b, S_j^h\right)$ is the utility from the racial composition of the neighborhood, with $S_j^b$ the share of neighborhood $j$ residents that are black and $S_j^h$ the share of neighborhood $j$ residents that are hispanic. We specify this function is quadratic in the black and hispanic shares:

$$f^{\tau}\left(S_j^b, S_j^h\right) \;=\; a_1^{\tau} S_j^b + a_2^{\tau}\left(S_j^b\right)^2 + a_3^{\tau} S_j^h + a_4^{\tau}\left(S_j^h\right)^2 + a_5^{\tau} S_j^b S_j^h \tag{9}$$

An obvious challenge in estimating the parameters in equations (8) and (9) is that $r_j$, $S_j^b$, and $S_j^h$ will be correlated with $A_j^{\tau}$ and $A_j^{\tau}$ is unobserved. For example, there may be neighborhoods with a relatively high value of $A_j^{\tau}$ for high-credit score, old, homeowning white households. Since $A_j^{\tau}$ is relatively high for these households, they will likely be the occupants. This implies the neighborhoods are likely to also have high house prices and low black- and hispanic- shares. To deal with this endogeneity we need an instrumental variables approach.

In estimation we will not be able to strictly identify preferences for racial shares in utility if certain amenities tend to be attached to certain racial shares. Continuing with the metaphor of the introduction, suppose that wherever black households live they plant pine trees and wherever white households live they plant spruce trees. In this case, we no longer directly estimate preferences for black households as neighbors, but rather the extent to which households have preferences for black neighbors *and* preferences for pine trees (relative to neighborhoods with white neighbors and spruce trees).

## 4.2 Discussion of our IV Approach

In estimation, we allow each type to have its own value of amenities in each neighborhood (Census tract), $A_j^{\tau}$. With this method, different types of households may sort themselves into neighborhoods simply because they value amenities of those neighborhoods differently; or, they may sort because of different valuations of rents and the demographic makeup of neighborhoods. Estimating parameters in this framework is straightforward: We uncover $\kappa^{\tau}$ and $\delta_j^{\tau}$ for all $j$ and $\tau$ via maximium likelihood, and then given a value of $a_r^{\tau}$ that we take from Davis, Gregory, Hartley, and Tan (2021) use instrumental variables to estimate all $a_1^{\tau}, \ldots, a_5^{\tau}$ for all $\tau$. Once these parameters are known, we compute $A_j^{\tau}$ as a residual.

Finding instruments to identify preferenes over the racial composition of neighborhoods is difficult. A few recent examples to identify these preferences include Card, Mas, and Rothstein (2008) who use a regression discontinuity approach, Almagro, Chyn, and Stuart (2022) who use random public housing demolitions and BLP-style instruments similar to Davis, Gregory, Hartley, and Tan (2021), and Caetano and Maheshri (2021) who use long lags of racial shares after controlling for inflows in a model-consistent way, just to name a few. Our approach, which we believe is new, is to use a Bartik-style instrument.[8]

Our instrument exploits variation across metro areas in the share of households that are black, hispanic, and white, much in the same way that the original Bartik instrument exploits variation across metro areas in the share of employment that is accounted for by different industries. Our instrument uses national data to predict where each type of household will locate inside each MSA; this is analogous to the Bartik instrument that assigns local predicted growth in each industry equal to national employment growth for that industry. Finally, given our type-by-type prediction on within-MSA location decisions is based only on national data, any variation (across MSAs) in predicted racial shares in each location can only be generated by differences in the distribution of types across MSAs. This is analogous to saying that in the original Bartik application, the predicted employment growth for each industry is constant, but predicted employment growth varies across MSAs because the mix of industries varies across MSAs.

In the original analysis of Bartik (1991), the instrument is used to forecast one statistic per MSA per time period, employment growth. In our application, we need to forecast the black share and the hispanic share in every Census tract in an MSA. Conceptually, forecasting these shares in multiple tracts in our application is equivalent to forecasting employment growth in multiple time periods in the original Bartik paper.

To explain how we construct our instrument, first note that the total population of Census tract $\ell$ in MSA $m$ can be written as $\sum_{\tau'} pop^{\tau'}_{\ell,m}$, where $pop^{\tau'}_{\ell,m}$ is the population of type $\tau'$ living in tract $\ell$ in MSA $m$. Thus the black share in tract $\ell$ in MSA $m$, $S^b_{\ell,m}$, can be written as[9]

$$S^b_{\ell,m} = \frac{\sum_{\tau} \mathcal{I}\left(\tau \in black\right) pop^{\tau}_{\ell,m}}{\sum_{\tau'} pop^{\tau'}_{\ell,m}}$$

Note that

$$pop^{\tau}_{\ell,m} = N^{\tau}_{m} \rho^{\tau}_{\ell,m}$$

where $N^{\tau}_{m}$ is the total number of type $\tau$ households living in MSA $m$ and $\rho^{\tau}_{\ell,m}$ is the probability that a type $\tau$ household living in MSA $m$ chooses to live in tract $\ell$. Now, make

[8] Those readers familiar with the first verison of this paper may recall that in that earlier work, we focused on the Qualifying Census Tract (QCT) designation of the Low Income Housing Tax Credit Program (LIHTC) as a valid instrument. After some analysis, we realized that we actually were using two instruments, a Bartik-style instrument we explain now and the QCT designation, and that all of the power in our first stage was from the Bartik-style instrument. This explains why we have dropped QCT from our analysis and have focused on the Bartik instrument.

[9] In what follows, the hispanic share is computed analogously.

that substitution and divide both the numerator and the denominator by the total MSA population:

$$S^b_{\ell,m} = \frac{\sum_{\tau} \mathcal{I}(\tau \in black)\, s^{\tau}_m \rho^{\tau}_{\ell,m}}{\sum_{\tau'} s^{\tau'}_m \rho^{\tau}_{\ell,m}} \tag{10}$$

In equation (10), $s^{\tau}_m$ is the share of the MSA population that is accounted for by type $\tau$ households.

To construct the instrument, we replace $\rho^{\tau}_{\ell,m}$, the actual probability a type $\tau$ household chooses tract $\ell$ in MSA $m$, with a predicted probability density *that only varies with the income percentile of the tract.*[10] Denote the income percentile associated with tract $\ell$ in MSA $m$ as $q(\ell, m)$ and denote the predicted probability density that specific tract is chosen by type $\tau$ as $\widehat{\rho}^{\tau}_{q(\ell,m)}$. Given this, our predicted black share in tract $\ell$ of MSA $m$ is

$$\widehat{S}^b_{\ell,m} = \frac{\sum_{\tau} \mathcal{I}(\tau \in black)\, s^{\tau}_m \widehat{\rho}^{\tau}_{q(\ell,m)}}{\sum_{\tau'} s^{\tau'}_m \widehat{\rho}^{\tau'}_{q(\ell,m)}} \tag{11}$$

To construct predicted probability densities, we regress the log of $\rho^{\tau}_{\ell,m}$ on MSA fixed effects[11] and a 7th order polynomial in the income percentile associated with tract $\ell$ in MSA $m$, $q(\ell, m)$. We run this regression separately for each type, but for each type we pool all tracts in all MSAs. Thus, for any type and any income percentile of a tract, the predicted probability density based on this regression does not vary across MSAs (except for the MSA fixed effects).

To give some intuition on how our Bartik-style IV works in practice, we construct a simple example. Suppose there are three types of households in our data – black, hispanic, and white – and that each MSA in our sample has exactly three tracts: low-income, middle-income, and high-income. In the first step, we use national data to estimate the probability that each type of households lives in one of the three tracts. Estimates from our data for each of the three types of households in each of the low- middle- and high-income tracts are shown in the top panel of table 1.

Now consider predicting the black, hispanic, and white shares of each of the three tracts in two of the metro areas in our sample. The first metro area shown in the middle panel, York-Hanover, PA, has a population that is 2.1% black, 2.7% hispanic, and 95.2% white;

[10] For notation reasons, we switch from probabilities to probability densities to handle the fact that different MSAs have different numbers of tracts.

[11] We include MSA fixed effects to account for the fact that MSAs vary in the total number of tracts, so MSAs with fewer tracts will by construction have higher choice probabilities in every tract.

Table 1: A Simple Example of the Bartik-Style Instrument

Probabilities over Locations using National Data

| | Black | Hispanic | White |
|---|---|---|---|
| Low Income Tract | 41.9% | 35.2% | 18.8% |
| Middle Income Tract | 32.5% | 34.4% | 35.0% |
| High Income Tract | 25.6% | 30.4% | 46.2% |

Racial Shares by Tract: York-Hanover, PA

| | Predicted | | | Actual | | |
|---|---|---|---|---|---|---|
| | Black | Hispanic | White | Black | Hispanic | White |
| Overall | 2.1% | 2.7% | 95.2% | 2.1% | 2.7% | 95.2% |
| By Tract: | | | | | | |
| - Low Income | 4.5% | 4.8% | 90.7% | 10.9% | 7.2% | 82.0% |
| - Middle Income | 2.0% | 2.7% | 95.4% | 3.6% | 2.3% | 94.2% |
| - High Income | 1.2% | 1.8% | 97.0% | 3.7% | 1.9% | 94.4% |

Racial Shares by Tract: Trenton, NJ

| | Predicted | | | Actual | | |
|---|---|---|---|---|---|---|
| | Black | Hispanic | White | Black | Hispanic | White |
| Overall | 8.4% | 17.5% | 74.1% | 8.4% | 17.5% | 74.1% |
| By Tract: | | | | | | |
| - Low Income | 14.9% | 26.1% | 59.0% | 38.3% | 17.7% | 44.0% |
| - Middle Income | 7.9% | 17.4% | 74.8% | 17.5% | 8.3% | 74.1% |
| - High Income | 5.2% | 12.8% | 82.1% | 6.5% | 5.7% | 87.8% |

and the second metro area (Trenton, NJ) shown in the bottom panel has a population that is 8.4% black, 17.5% hispanic, and 74.1% white.[12] Given these overall metro-area type shares, first three columns of the middle and bottom panels show the predicted racial shares and, for comparison, the final three columns show the actual racial shares. By construction, the variation in predicted shares at the Census tract level is driven only by variation in metro-area racial shares. As the table illustrates, both the predicted and actual racial shares vary quite considerably, and the predicted racial shares seem to be correlated with the actual. For example, in York-Hanover, PA, the predicted black share of the low-income tract is 4.5% (actual is 10.9%), whereas in Trenton, NJ, the predicted black share of the low-income tract is 14.9% (actual is 38.3%). In both metro areas, the share of black households in the lowest-income tracts is higher than predicted, but the instrument exploits the fact that Trenton, NJ has more black households than York-Hanover, PA to predict that the share of black

[12]York-Hanover, PA and Trenton, NJ are about 120 miles apart.

households in low-income tracts is higher in Trenton than in York.

## 4.3 Discussion of Identification

### 4.3.1 Rotemberg Weights

To understand the source of variation in our Bartik-style instrument, we follow Goldsmith-Pinkham, Sorkin, and Swift (2020), hereafter GSS, and plot the Rotemberg weights of our instrument. For the purposes of computing Rotemberg weights, we consider 5,400 instruments: 54 types by 100 income percentiles in each MSA. The Rotemberg weights identify the fraction of variation in the predicted black (hispanic) share that comes from each of the instruments.

One reason we describe our predicted black (hispanic) share as resulting from a "Bartik-style" instrument is that our predicted share is actually the ratio of two Bartiks. The denominator is the total fraction of all households in the MSA that choose the tract and the numerator is the total fraction of all households in the MSA that choose the tract and are black: See equation (11). Our predicted black (hispanic) share is therefore a non-linear function of the type shares, $s_m^\tau$, and we cannot apply a direct application of the procedure described in GSS to compute Rotemberg weights. Instead, we take a linear approximation of our predicted black (hispanic) share and compute the Rotemberg weights of this linear approximation.

In what follows, we show our method to compute Rotemberg weights for predicted black shares.[13] As before, define the actual predicted black share at income quantile $q$ associated with tract $\ell$ in MSA $m$ as

$$\widehat{S}_{l,m}^{b} = \frac{\sum_{\tau} I\left(\tau \in black\right) s_m^\tau \hat{\rho}_{q(\ell,m)}^\tau}{\sum_{\tau'} s_m^{\tau'} \hat{\rho}_{q(\ell,m)}^\tau}$$

To conserve on notation, we will replace both $q\left(\ell,m\right)$ and $\ell$ with $q$ in what follows. Note

[13] The method is identical for hispanic shares.

that the derivative of $\widehat{S}^b_{q,m}$ with respect to $s^\tau_m$ is equal to

$$\begin{aligned}\frac{\partial \widehat{S}^b_{q,m}}{\partial s^\tau_m} &= \frac{I\left(\tau = black\right)\hat{\rho}^\tau_q}{\sum\limits_{\tau'} s^{\tau'}_m \hat{\rho}^{\tau'}_q} - \frac{\hat{\rho}^\tau_q \sum\limits_{\tau''} I\left(\tau'' = black\right) s^\tau_m \hat{\rho}^{\tau''}_q}{\left(\sum\limits_{\tau'} s^{\tau'}_m \hat{\rho}^{\tau'}_q\right)^2} \\ &= \left(\frac{\hat{\rho}^\tau_q}{\sum\limits_{\tau'} s^{\tau'}_m \hat{\rho}^{\tau'}_q}\right)\left[I\left(\tau = black\right) - \widehat{S}^b_{q,m}\right]\end{aligned}$$

Continuing, define $\bar{s}^\tau$ as the national average share of type $\tau$ in an MSA, and define $\bar{S}^b_q$ as the predicted black share in income quantile $q$ when all type shares in an MSA are equal to their national average, i.e.

$$\bar{S}^b_q \equiv \frac{\sum\limits_{\tau} I\left(\tau \in black\right) \bar{s}^\tau \hat{\rho}^\tau_q}{\sum\limits_{\tau'} \bar{s}^{\tau'} \hat{\rho}^{\tau'}_q}$$

Given this notation, a first-order Taylor series approximation of $\widehat{S}^b_{q,m}$ around $\bar{S}^b_q$ is equal to

$$\begin{aligned}\widehat{S}^b_{q,m} &\approx \bar{S}^b_q + \sum_\tau \left.\frac{\partial \widehat{S}^b_{q,m}}{\partial s^\tau_m}\right|_{\bar{s}^\tau \forall \tau} \left(s^\tau_m - \bar{s}^\tau\right) \\ \text{where} \quad \left.\frac{\partial \widehat{S}^b_{q,m}}{\partial s^\tau_m}\right|_{\bar{s}^\tau \forall \tau} &= \left(\frac{\hat{\rho}^\tau_q}{\overline{pop}_q}\right)\left[I\left(\tau = black\right) - \bar{S}^b_q\right]\end{aligned}$$

$\overline{pop}_q$ is the population of the tract with income percentile $q$ when all type shares are equal to the national average: $\overline{pop}_q = \sum\limits_{\tau'} \bar{s}^{\tau'} \hat{\rho}^{\tau'}_q$. After simplification, the above reduces to

$$\begin{aligned}\widehat{S}^b_{q,m} &\approx \bar{S}^b_q + \sum_\tau g^\tau_q s^\tau_m \\ \text{where} \quad g^\tau_q &= \left(\frac{\hat{\rho}^\tau_q}{\overline{pop}_q}\right)\left[I\left(\tau = black\right) - \bar{S}^b_q\right]\end{aligned} \tag{12}$$

Thus we can approximate $\widehat{S}^b_{q,m}$ as a linear function of type shares. To check the quality of this approximation, we regress $\widehat{S}^b_{q,m}$ on the linear approximation as given by equation (12) for all 5,400 instruments and get a slope of 0.90 and an $R^2$ of 0.99.[14]

Given this linear approximation, we compute Rotemberg weights as follows. First, we regress the actual predicted $\widehat{S}^b_{q,m}$ on a full set of MSA dummy variables and a 7th order polynomial in income percentile. Define the residual from this regression evaluated for MSA

[14] The results for the hispanic share are similar.

$m$ at quantile $q$ as $\widehat{S}^{b\perp}_{q,m}$. Assuming there are $M$ MSAs in the sample, define $\widehat{S}^{b\perp}_{q}$ as the $Mx1$ vector of residuals evaluated at quantile $q$

$$\widehat{S}^{b\perp}_{q} \;=\; \left[\widehat{S}^{b\perp}_{q,1} \quad \widehat{S}^{b\perp}_{q,2} \quad \ldots \quad \widehat{S}^{b\perp}_{q,M}\right]'$$

Similarly, define $s^{\tau}$ as the $Mx1$ vector of type $\tau$ shares across all MSAs:

$$s^{\tau} \;=\; \left[s^{\tau}_{1} \quad s^{\tau}_{2} \quad \ldots \quad s^{\tau}_{M}\right]'$$

Then for any type $\hat{\tau}$ and income percentile $\hat{q}$ we define its Rotemberg weight as[15]

$$\frac{g^{\hat{\tau}}_{\hat{q}} \;\left[s^{\hat{\tau}}\right]' \cdot \widehat{S}^{b\perp}_{\hat{q}}}{\sum_{q}\sum_{\tau} \; g^{\tau}_{q} \;\left[s^{\tau}\right]' \cdot \widehat{S}^{b\perp}_{q}}$$

where $g^{\tau}_{q}$ is as defined in equation (12) for all $q$ and $\tau$.

The top panel of table 2 summarizes the Rotemberg weights for predicting the black share. Black households account for 84.5% of these Rotemberg weights. Hispanic households account for 12.0% and white households account for only 3.5%. About 40% (740/1,800) of the Rotemberg weights for predicting the black share of neighborhoods attributable to white households are negative, but the sum of these negative Rotemberg weights is quite small at -1.4%; the other, positive Rotemberg weights for white households sums to 4.9%. The bottom two rows of this panel show that the Rotember weights for predicting the black share for black renting households is 65.7% and for black home-owning households it is 18.9%.

The bottom panel of table 2 summarizes the Rotemberg weights for predicting the Hispanic share. All Rotemberg weights for predicting the Hispanic share are nonnegative. Hispanic households account for 88.3% of these weights, with white households accounting for 7.1% and black households accounting for 4.6%. The bottom two rows of this panel show that the Rotember weights for predicting the hispanic share for hispanic renting households is 67.3% and for hispanic home-owning households it is 21.1%. Overall, both panels of table 2 show that the Rotemberg weights of renting households for predicting own-race racial shares sum to about 2/3rds.

The top panel of Figure 1 shows the Rotemberg weights of black renting households for predicting the black share.[16] To keep the graph clean, we show the sum of the weights in

[15] Note that when we estimate preferences over black and hispanic shares, we need to generate a value of $\widehat{S}^{b}_{q(\ell,m),m}$ appropriate for each tract in each MSA. For the purposes of understanding the source of variation in the predicted black and hispanic shares, that is not necessary: it is acceptable to divide all tracts in an MSA into a fixed number of quantiles based on tract income.

[16] We do not graph the Rotemberg weights for black homeowning households because these weights are

Table 2: Distribution of Rotemberg Weights

Predicted Black Share

| Race | Owner/Renter | Sum of Rotemberg Weights |
|---|---|---|
| Black | | 0.845 |
| Hispanic | | 0.120 |
| White* | | 0.035 |
| Black | Renter | 0.657 |
| Black | Owner | 0.189 |

Predicted Hispanic Share

| Race | Owner/Renter | Sum of Rotemberg Weights |
|---|---|---|
| Black | | 0.046 |
| Hispanic | | 0.883 |
| White | | 0.071 |
| Hispanic | Renter | 0.673 |
| Hispanic | Owner | 0.211 |

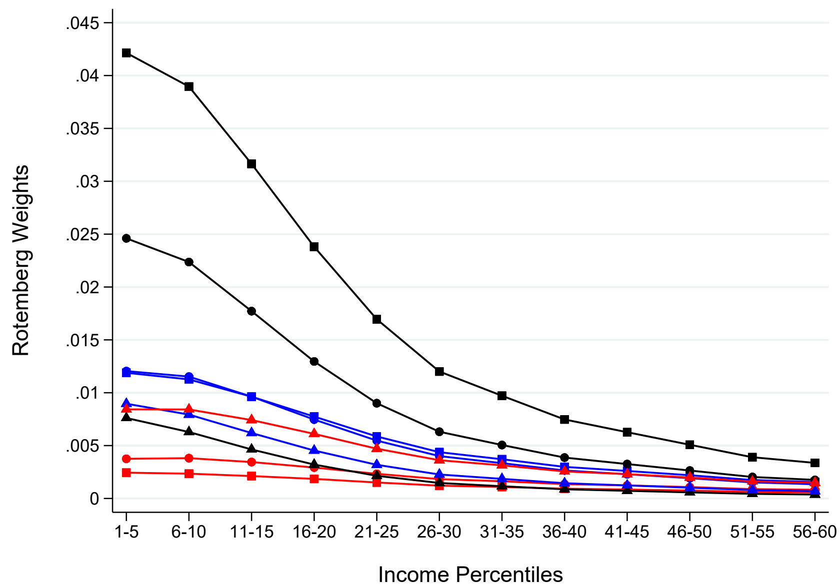
* White types have 740 (out of 1800) negative Rotemberg weights for predicted black shares. The average weight when negative is $-1.86 \times 10^{-5}$ and the sum of these negative weights is -0.014.

5 percentile income bins, i.e. income percentiles 1-5, 6-10, and so forth. We do not show results between the 61st and 100th income percentiles since there is minimal variation in that range. There are 9 lines on this graph, one for each type of black renting household. The different colors correspond to different credit bins – black for lowest, blue for middle, and red for highest – and the different markers refer to different ages – square for youngest, circle for middle aged, and triangle for oldest. The figure shows that a disproportionate amount of variation in predicted black shares is accounted for by young- and middle-aged black renting households with low credit scores locating in lower-income tracts, the black lines with square and circle markers.

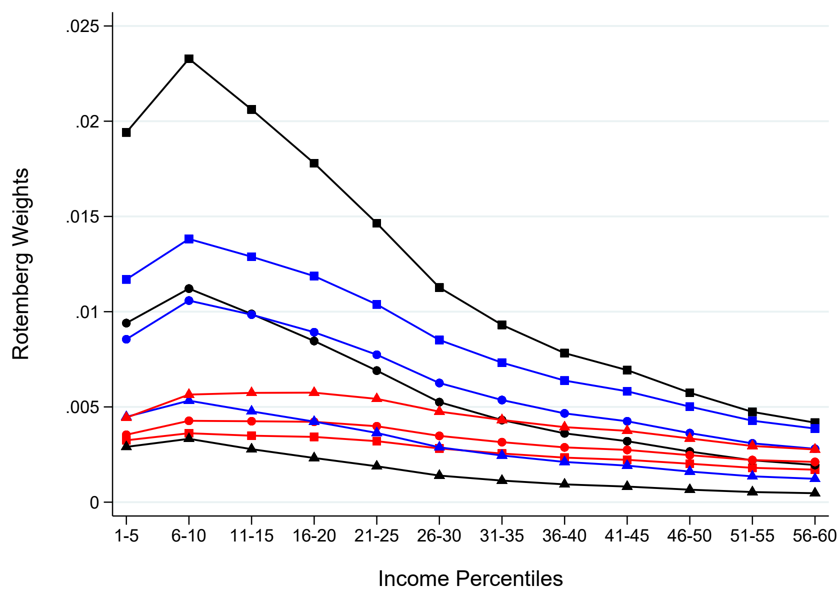
The bottom panel of Figure 1 shows the Rotember weights for predicting the hispanic share for hispanic renting households. The formatting of the bottom panel is identical to the top panel. The panel shows that four types of Hispanic households account for a disproportionate amount of variation in predicted hispanic shares: young- and middle-aged black renting households with low and middle-tier credit scores, the black and the blue lines with square and circle markers.

relatively very small.

Figure 1: Rotemberg Weights




(a) Weights for Predicted Black Share, Black Renters




(b) Weights for Predicted Hispanic Share, Hispanic Renters

Notes: black line is lowest credit score, blue line is medium credit score, red line is highest credit score, square marker is youngest age, circle marker is middle age, and triangle marker is oldest age.

Table 3: Balance Test Results

Panel A: Predicted Black Share

| Outcome | mean | sd | Coefficient on $\widehat{S}_j^b$ IV | SE | IV2 | SE | p-values IV | IV2 |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Distance to river (mi) | 3.316 | 3.562 | 0.712 | (1.105) | 1.321 | (1.649) | 0.520 | 0.423 |
| Fraction flat planes | 0.410 | 0.423 | -0.167 | (0.113) | -0.174 | (0.102) | 0.141 | 0.089 |
| Public transit < .25 miles | 0.023 | 0.103 | 0.053 | (0.057) | 0.033 | (0.048) | 0.354 | 0.495 |
| Public transit < .5 miles | 0.059 | 0.189 | 0.085 | (0.094) | 0.054 | (0.084) | 0.369 | 0.522 |
| Road network density | 2.895 | 6.165 | 0.675 | (1.402) | 0.885 | (1.644) | 0.630 | 0.591 |
| $S_j^b$ | 0.156 | 0.250 | 1.221 | (0.021) | 1.226 | (0.024) | 0.000 | 0.000 |
| $S_j^h$ | 0.112 | 0.179 | -0.426 | (0.058) | -0.411 | (0.054) | 0.000 | 0.000 |

Panel B: Predicted Hispanic Share

| Outcome | mean | sd | Coefficient on $\widehat{S}_j^h$ IV | SE | IV2 | SE | p-values IV | IV2 |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Distance to river | 3.316 | 3.562 | 1.029 | (1.761) | 2.015 | (1.994) | 0.559 | 0.312 |
| Fraction flat planes | 0.410 | 0.423 | 0.517 | (0.199) | 0.246 | (0.235) | 0.009 | 0.294 |
| Public transit < .25 miles | 0.023 | 0.103 | -0.017 | (0.063) | -0.091 | (0.066) | 0.786 | 0.163 |
| Public transit < .5 miles | 0.059 | 0.189 | 0.041 | (0.120) | -0.116 | (0.130) | 0.731 | 0.374 |
| Road network density | 2.895 | 6.165 | -3.845 | (2.089) | -5.097 | (2.863) | 0.066 | 0.075 |
| $S_j^b$ | 0.156 | 0.250 | -1.518 | (0.177) | -1.397 | (0.256) | 0.000 | 0.000 |
| $S_j^h$ | 0.112 | 0.179 | 1.383 | (0.022) | 1.373 | (0.034) | 0.000 | 0.000 |

### 4.3.2 Balance Tests

For our estimates to be unbiased, the instruments must be uncorrelated with amenities: The following must hold for all $\ell$ and $\tau$ for the black share

$$\text{cov}\left(A^{\tau\perp}_{\ell,m}, \widehat{S}^{b\perp}_{\ell,m}\right) = 0 \tag{13}$$

with the analogous expression holding for hispanic share. In the above equation, $A^{\tau\perp}_{\ell,m}$ would be the residual of a regression of $A^{\tau}_{\ell,m}$ – if it were observable – on a full set of MSA dummy variables and a 7th order polynomial in income percentile.

We wish to test equation (13), but amenities are not directly observable. Instead we try to find variables that may proxy for amenities, and test if equation (13) holds for these proxies. To the extent that certain amenities are associated with certain racial groups (i.e. pine trees and black households), and our estimates for preferences for the black and hispanic shares of neighborhoods bundle both preferences for race and amenities associated with race, we do not want to include proxies for $A^{\tau\perp}_{\ell,m}$ that may be influenced by household or government choices potentially related to racial shares.

For proxies for amenities, we use tract distance to the nearest river (in miles),[17] the fraction of the tract that is a flat plain, a dummy variable if the tract is within 0.25 miles of public transit, a dummy variable if the tract is within 0.50) miles of public transit, and an estimate of the road network density of the tract. The top panel of Table 3 shows results of these balance tests for predicted black shares and the bottom panel shows results for predicted hispanic shares. Columns (1) and (2) of Table 3 show the means and standard deviations of these variables. Columns (3) and (5) show estimates from two different regressions of the outcome variable (i.e. distance to river) on our instrument after controlling for MSA fixed effects and a 7th order polynomial in income; columns (4) and (6) show the standard errors; and columns (7) and (8) show p-values for the null hypothesis that the coefficient is zero. The column marked "IV" is for our actual instrument and the column marked "IV2" is the linear approximation of our instrument from equation (12), but only including black or hispanic households that rent, the source of most of the variation of the instrument according to the Rotemberg weights. Columns (4), (6), (7) and (8), show that we systematically fail to reject the null that our instrument is uncorrelated with these variables.

Columns (3) - (8) of the last two rows of the top and bottom panels show the results from regressing actual black or hispanic shares on either predicted black share (top panel) or predicted hispanic share (bottom panel). As before, these regressions also include MSA fixed effects and control for tract income percentile.[18] Shown in the top panel, the predicted black share strongly positively predicts the actual black share and strongly negatively predicts the actual hispanic share. Conversely, the bottom panel shows the predicted hispanic share strongly negatively predicts the actual black share and strongly positively predicts the actual hispanic share.

## 4.4 Discussion of Estimates

Table 4 provides a summary of our estimates. Column (1) of shows the type index and (2) reports the percentage of the estimation sample accounted for by that type. Columns (3)-(6) show the race, age (y=young, m=middle-aged, and o=old), homewonership tenure (r=rent, o=own), and credit score bin (l=low, m=middle, h=high) of the type. Column (7) reports the average share of black households in the Census tracts in which that type tends to live and column (8) shows the average derivative of utility that type would experience from an increase in the share of black households in the Census tracts in which that type tends

[17]We require that some tract in the CBSA lie within 5 miles of a river for any tract in that CBSA to be included in our analysis.

[18]These regressions are not our actual first stages, but we find the results in this table to be informative on the relevance of the instruments.

# Table 4: Summary of Estimates of Prferences over Race

| Type | Sample % | Race | Age | Tenure | Credit | Avg $S_j^b$ | Avg $\Delta\delta_j/\Delta S_j^b$ | Avg $S_j^h$ | Avg $\Delta\delta_j/\Delta S_j^h$ | $\delta_j^{95}-\delta_j^5$ |
|---|---|---|---|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| 1 | 2.2% | | y | r | l | 0.47 | 1.22 | 0.10 | 2.32 | 1.54 |
| 2 | 1.2% | | y | r | m | 0.37 | 1.11 | 0.11 | 1.31 | 1.06 |
| 3 | 0.6% | | y | r | h | 0.26 | 0.41 | 0.10 | -0.01 | 0.47 |
| 4 | 0.4% | | y | o | l | 0.38 | 1.21 | 0.09 | 1.16 | 1.19 |
| 5 | 0.5% | | y | o | m | 0.28 | 0.59 | 0.09 | 0.96 | 0.58 |
| 6 | 0.6% | | y | o | h | 0.17 | -0.21 | 0.08 | 0.37 | 0.94 |
| 7 | 1.1% | | m | r | l | 0.50 | 0.89 | 0.09 | 1.64 | 1.36 |
| 8 | 0.9% | | m | r | m | 0.43 | 0.89 | 0.10 | 0.46 | 1.05 |
| 9 | 0.6% | Black | m | r | h | 0.31 | 0.42 | 0.09 | -0.51 | 0.35 |
| 10 | 0.4% | | m | o | l | 0.47 | 0.69 | 0.08 | -0.19 | 1.04 |
| 11 | 0.6% | | m | o | m | 0.38 | 0.48 | 0.08 | -0.42 | 0.58 |
| 12 | 0.9% | | m | o | h | 0.23 | 0.38 | 0.08 | -0.42 | 0.53 |
| 13 | 0.3% | | o | r | l | 0.56 | 0.60 | 0.08 | 0.90 | 1.46 |
| 14 | 0.5% | | o | r | m | 0.52 | 0.58 | 0.08 | 0.04 | 1.10 |
| 15 | 1.0% | | o | r | h | 0.39 | 0.25 | 0.07 | -0.71 | 0.34 |
| 16 | 0.1% | | o | o | l | 0.58 | 1.03 | 0.07 | 0.31 | 1.81 |
| 17 | 0.2% | | o | o | m | 0.51 | 0.88 | 0.07 | 0.19 | 1.33 |
| 18 | 0.4% | | o | o | h | 0.33 | 0.42 | 0.07 | -0.47 | 0.41 |
| Sum | 12.3% | | | | Avg | 0.39 | 0.73 | 0.09 | 0.70 | 0.97 |
| 19 | 1.6% | | y | r | l | 0.14 | 1.61 | 0.37 | 1.55 | 1.42 |
| 20 | 1.4% | | y | r | m | 0.11 | 0.78 | 0.36 | 1.49 | 1.40 |
| 21 | 0.8% | | y | r | h | 0.09 | -0.85 | 0.29 | 1.11 | 1.74 |
| 22 | 0.3% | | y | o | l | 0.12 | 1.21 | 0.33 | 1.35 | 1.23 |
| 23 | 0.5% | | y | o | m | 0.10 | -0.28 | 0.31 | 1.29 | 1.70 |
| 24 | 0.7% | | y | o | h | 0.07 | -1.82 | 0.22 | 0.82 | 1.82 |
| 25 | 0.7% | | m | r | l | 0.14 | 1.50 | 0.37 | 1.09 | 1.10 |
| 26 | 0.9% | | m | r | m | 0.11 | 0.88 | 0.38 | 0.93 | 0.92 |
| 27 | 0.8% | Hisp | m | r | h | 0.09 | -0.37 | 0.32 | 0.93 | 1.25 |
| 28 | 0.3% | | m | o | l | 0.13 | 0.72 | 0.35 | 0.79 | 0.76 |
| 29 | 0.6% | | m | o | m | 0.10 | 0.27 | 0.32 | 0.77 | 1.10 |
| 30 | 1.1% | | m | o | h | 0.07 | -0.30 | 0.23 | 0.58 | 1.34 |
| 31 | 0.1% | | o | r | l | 0.14 | 1.43 | 0.41 | 1.13 | 1.18 |
| 32 | 0.3% | | o | r | m | 0.11 | 0.74 | 0.40 | 0.99 | 0.94 |
| 33 | 1.0% | | o | r | h | 0.08 | -0.47 | 0.31 | 0.65 | 1.08 |
| 34 | 0.0% | | o | o | l | 0.14 | 1.94 | 0.38 | 1.67 | 1.38 |
| 35 | 0.1% | | o | o | m | 0.12 | 0.54 | 0.36 | 1.04 | 0.96 |
| 36 | 0.4% | | o | o | h | 0.08 | -0.26 | 0.25 | 1.16 | 1.19 |
| Sum | 11.6% | | | | Avg | 0.10 | 0.31 | 0.33 | 1.07 | 1.30 |
| 37 | 5.6% | | y | r | l | 0.13 | 1.05 | 0.11 | 1.03 | 0.81 |
| 38 | 6.1% | | y | r | m | 0.09 | 0.50 | 0.10 | 0.87 | 1.40 |
| 39 | 5.4% | | y | r | h | 0.07 | -1.19 | 0.08 | -0.23 | 2.23 |
| 40 | 1.5% | | y | o | l | 0.10 | 1.03 | 0.09 | 0.65 | 1.29 |
| 41 | 3.3% | | y | o | m | 0.08 | -0.86 | 0.08 | 0.57 | 2.07 |
| 42 | 6.8% | | y | o | h | 0.06 | -2.90 | 0.06 | 0.08 | 2.67 |
| 43 | 2.7% | | m | r | l | 0.13 | 1.12 | 0.11 | 0.52 | 0.55 |
| 44 | 4.1% | | m | r | m | 0.09 | 0.87 | 0.10 | 0.25 | 0.92 |
| 45 | 6.5% | White | m | r | h | 0.06 | -0.22 | 0.07 | -0.35 | 1.63 |
| 46 | 1.3% | | m | o | l | 0.10 | 1.04 | 0.08 | -0.11 | 0.94 |
| 47 | 3.7% | | m | o | m | 0.08 | 0.76 | 0.08 | -0.18 | 1.44 |
| 48 | 11.5% | | m | o | h | 0.05 | -0.31 | 0.06 | -0.22 | 2.04 |
| 49 | 0.5% | | o | r | l | 0.13 | 1.10 | 0.11 | 0.41 | 0.49 |
| 50 | 1.8% | | o | r | m | 0.09 | 0.86 | 0.09 | 0.20 | 0.67 |
| 51 | 10.5% | | o | r | h | 0.06 | -0.27 | 0.07 | -0.51 | 1.49 |
| 52 | 0.2% | | o | o | l | 0.12 | 1.05 | 0.09 | -0.01 | 0.73 |
| 53 | 0.7% | | o | o | m | 0.09 | 1.04 | 0.08 | -0.10 | 0.95 |
| 54 | 4.0% | | o | o | h | 0.06 | 0.26 | 0.06 | -0.37 | 1.75 |
| Sum | 76.1% | | | | Avg | 0.08 | -0.15 | 0.08 | 0.05 | 1.61 |

For age: y = young, m = middle-aged, o = old. For tenure: r = renter, o = owner. For credit: l = low, m = middle, h = high.

to live. Similarly, column (9) reports the average share of hispanic households in the Census tracts in which which that type tends to live and column (10) shows the average derivative of utility that type would experience from an increase in the share of hispanic households in the Census tracts in which that type tends to live. Note that the values reported in columns (7)-(10) are computed as weighted averages over all tracts in which the type may live, with the weights being the probability that the type lives in the tract.[19]

The top panel of the table show results for black types, the middle panel show results for hispanic types, and the bottom panel show results for white types. Focusing on the bottom row of each of the panels panel, black households account for 12.3% of our sample, hispanic households account for 11.6% of our sample, and white households account for 76.1% of our sample. Table 4 shows that same-race sorting is a prominent feature of our data. Columns (7) and (9) show that, on average, black households live in Census tracts that are 39% black, hispanic households live in Census tracts that are 33% black and white households live in Census tracts that are comprised of 84% white households.

Columns (8) and (10) show the derivative of utility with respect to exogenous changes in the tract-level black share (8) or hispanic share (10). Shown in the bottom row of the top and middle panels, on average both black and hispanic households receive additional utility from an increase in black and hispanic shares. The bottom panel shows that white households are roughly indifferent to an increase in hispanic shares and, on average and with considerable heterogeneity, white households experience disutility from an exogenous increase in the share of black households living in their Census tracts. The white type experiencing the largest disutility are young, homeowning, high-credit score households, type 42, accounting for 6.8% of our sample. The derivative of utility of this type with respect to the black share of the population is -2.90, such that if the black share increases by one percentage point, utility falls by -0.029. For comparison, a twenty percent increase in rental prices generates approximately the same decline in utility.

Finally, column (11) illustrates the importance of racial preferences in accounting for location choice in our data. For column 11, we set $a_r^\tau = 0$ and $A_j^\tau = 0$ for all $\tau$ and all $j$ and then evaluate the the level of utility for each type in each tract; in this calculation, differences in black and hispanic shares entirely determine differences in utility across tracts. For each type, we sort tracts by the level of utility the tract provides; we then report in column (11) the level of utility for the type at the location representing the 95th percentile less the level of utility at the location representing the 5th percentile. These utility differentials

[19] For example, suppose there are two tracts A and B; and, thinking about column 8, suppose a particular type experiences a -1.0 derivative to utility with respect to the black share in tract A and a +1.0 derivative to utility with respect to the black share in tract B. If the probability that type lives in tract A is 0.20, then we would report a value in column 8 for that type of 0.6 which we compute as $0.2\,(-1.0) + 0.8\,(1.0)$.

attributable entirely to differences in racial composition across neighborhoods are huge: 0.97 for black households, 1.30 for hispanic households, and 1.61 for white households. On average, there is essentially no change to rent that can compensate types sufficiently to induce households to move from neighborhoods with their most desired demographic composition to neighborhoods with their least desired demographic composition.

**Figure XXX graphs how our estimates for YYY different household types maps to utility for various racial configurations.**

**Do we need a variance decomposition of racial shares vs amenities in determining where people live?**

# 5 Discussion of Demographic Stability

We now discuss if the demographic composition of neighborhoods is stable. We begin by introducing some notation and defining what we mean by stability. Ultimately, stability depends on the process by which expectations of neighborhood composition change.

For a given MSA $m$ with $J$ total tracts, denote $\mathcal{T}$ as $2J \times 1$ vector comprised of starting values of expectations of racial shares, $E\left[S_j^b\right]$ and $E\left[S_j^h\right]$ for all tracts. Let $g(\mathcal{T})$ be an expectations-generating function produced by our model that takes as a starting input $\mathcal{T}$ and produces a different vector of expectations $\mathcal{T}'$,

$$\mathcal{T}' = g(\mathcal{T}).$$

We define a steady state of $g$ as a vector of expectations $\mathcal{T}^*$ that generates, via $g$, an identical set of expectations, i.e.

$$\mathcal{T}^* = g(\mathcal{T}^*).$$

Before describing how $g(\mathcal{T})$ works, we now define a steady state that is consistent with our current data for each metro area. We start with the distribution of types by tract implied by our estimation sample and then simulate the model for 5 periods, our "burn in" period. During these 5 periods, we assume each household's type stays fixed. During the burn in period, we hold $\delta_j^\tau$ fixed for all types and all tracts. We use a 5-period burn in to ensure all types populate all tracts in our baseline steady state implied by our data. In a sense, the burn-in period smooths through sampling variability that may be in the data. After the burn-in, we use the resulting distribution of types by tract to compute our baseline vector for $\mathcal{T}$, $E\left[S_j^b\right] = S_j^b$ and $E\left[S_j^h\right] = S_j^h$ for all $j$.

Next, we compute the distribution of types across all tracts that results after running the decision model for one period such that all location choices are made and all types probabilistically evolve. For each tract, we compute the required additions ("births") or subtractions ("deaths") of the population of each type such that the resulting meausres of household types in each tract after all decisions are made and all types have stochastically evolved is constant in all tracts. The addition of type-specific births and deaths to each tract the model guarantees that the model-predicted distribution of types across tracts is stable and the vector $\mathcal{T}^*$ consistent with our data is a steady state. That is, the decisions implied by the model are consistent with expectations households have over racial shares and rental prices in each tract.

We now describe the $g(\mathcal{T})$ function that predicts how expectations evolve given any starting set of expectations $\mathcal{T}$. To start, denote the total number of households and the rental price in each tract in the data as $\mathcal{H}_j$ and $r_j$, respectively. Then, we compute $g(\mathcal{T})$ as follows:

1. Denote the guess of new rental prices $r'_j$.

2. Using equations (8) and (9), adjust $\delta_j^\tau$ appropriately for all $j$ and $\tau$ given the values of $E\left[S_j^b\right]$ and $E\left[S_j^h\right]$ from $\mathcal{T}$ and the guess $r'_j$, holding exogenous amenities $A_j^\tau$ fixed. Households assume this new value of $\delta_j^\tau$ is fixed forever when making decisions.

3. Simulate the model 99 periods and then compute new housing demand in each tract, $\mathcal{H}'_j$.

4. Update the guess of rental prices and repeat steps 2-3 until rental prices in each tract clear markets to satisfy

$$\log \mathcal{H}'_j - \log \mathcal{H}_j \quad = \quad \psi_j \left[\log r'_j - \log r_j\right]$$

   The housing supply elasticity in each tract $j$, $\psi_j$, is given by the estimates in Baum-Snow and Han (2022) with a floor elasticity of 0.025.[20]

5. Once we know rental prices $r'_j$ that clear housing markets given values of $E\left[S_j^b\right]$ and $E\left[S_j^h\right]$ from $\mathcal{T}$, compute simulated black and hispanic shares in each tract and call these $S_j^{b\prime}$ and $S_j^{h\prime}$.

6. Set the elements of $\mathcal{T}'$ equal to $S_j^{b\prime}$ and $S_j^{h\prime}$.

[20]In a handful of tracts, Baum-Snow and Han (2022) estimate a negative supply elasticity.

Given our procedure to compute $g(\mathcal{T})$, we test the stability of the steady state implied by the data by computing the eigenvalues and eigenvectors of the model at the steady state. To see why this is useful, suppose we perturb expectations of racial shares at the steady state – call these perturbed expectations as $\mathcal{T}'$ – and then measure how expectations evolve from this perturbed starting point, i.e. $T'' = g(\mathcal{T}')$. We can do this with a first-order linear approximation:

$$g(\mathcal{T}') - g(T^{*}) \approx \mathcal{G} \cdot [\mathcal{T}' - \mathcal{T}^{*}]$$

where $\mathcal{G}$ is a $2J$ by $2J$ vector of derivatives of $g$ evaluated at $\mathcal{T}^{*}$. Once we make appropriate substitutions, we get

$$[\mathcal{T}'' - \mathcal{T}^{*}] \approx \mathcal{G} \cdot [\mathcal{T}' - \mathcal{T}^{*}]$$

We compute the elements of $\mathcal{G}$ at $\mathcal{T}^{*}$ using numerical derivatives. Specifically, define $\widetilde{\mathcal{T}}_i^{*}$ as equal to $\mathcal{T}^{*}$ in all elements except for the $i^{th}$ element which we perturb by $\Delta_i$ units.[21] We set the $i^{th}$ column of $\mathcal{G}$ equal to $\left[g\left(\widetilde{\mathcal{T}}_i^{*}\right) - \mathcal{T}^{*}\right]/\Delta_i$. We repeat this computation for all $i = 1, \ldots, 2J$ elements of $\mathcal{T}^{*}$ to populate all the columns of $\mathcal{G}$.

Once we have estimates of $\mathcal{G}$, we compute its eigenvalues to determine whether the expectations of racial shares move away from or return to the steady state expectations implied by the data in response to a tiny perturbation to expectations. In other words, we ask if the system predicts expectations return to $\mathcal{T}^{*}$ if we start our model using expectations that are nearly but not exactly identical to $\mathcal{T}^{*}$. If all the eigenvalues of $\mathcal{G}$ are less than 1, the expectations converge back to the steady state; if at least one eigenvalue is greater than 1, expectations do not converge back to the starting point and in this instance, we say the steady state implied by the data is not stable.

Appendix table A.1 lists our full set of MSAs and results. Column (1) lists the MSA name, (2) lists the population, (3) lists the number of tracts, (4) lists the black share of the MSA population, and (5) lists the hispanic share of the population. Column (6) lists the share of eigenvalues of $\mathcal{G}$ of that MSA that have a value greater than 1. The results shown in column (6) very strongly suggest that the demographic composition of neighborhoods in the United States is inherently unstable – even using a model framework engineered to have a steady state at the current data. Only one MSA out of 141 in the sample, Rockingham County - Strafford County, NH, has zero eigenvalues greater than 1. Every other MSA has at least one eigenvalue greater than 1, and in fact the median MSA has 47% of its eigenvalues

[21] For each element $i$, we set $\Delta_i$ equal to $1.0 \times 10^{-6}$.

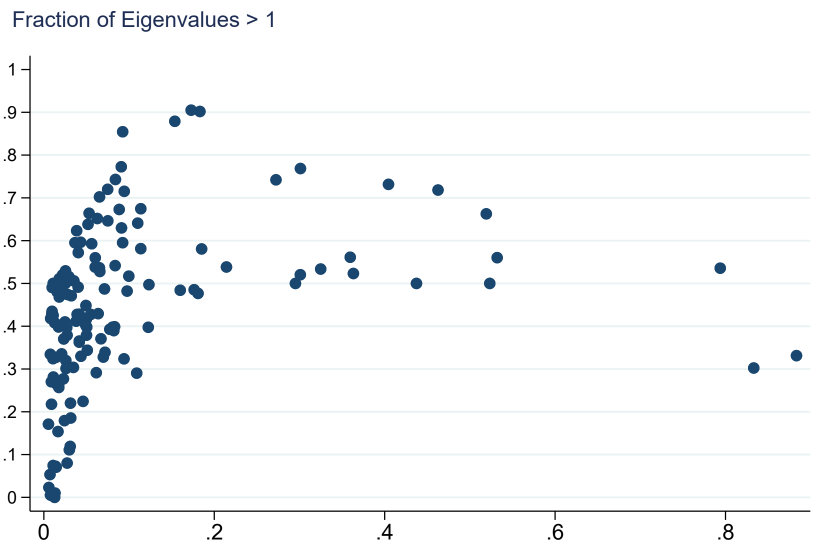
greater than 1 (shown at the end of the table).

Ultimately, the reason that the system is not stable is that households have very strong preferences over the racial composition of their neighbors. Rockingham County - Strafford County, NH, is stable because the population is almost entirely white; the black share is 1.3% and the hispanic share is 0.8%. The other MSAs that have a minimum of eigenvalues larger than 1 are also almost entirely white as well, for example Barnstable Town, MA (1.3% black and 2.3% hispanic), Duluth MN-WI (0.8% white and 1.0% hispanic), and Portland - South Portland, ME (0.9% white and 0.9% hispanic). The racial composition of neighborhoods in these metro areas is stable because the population is nearly entirely white and arithmetically stability is nearly guaranteed.

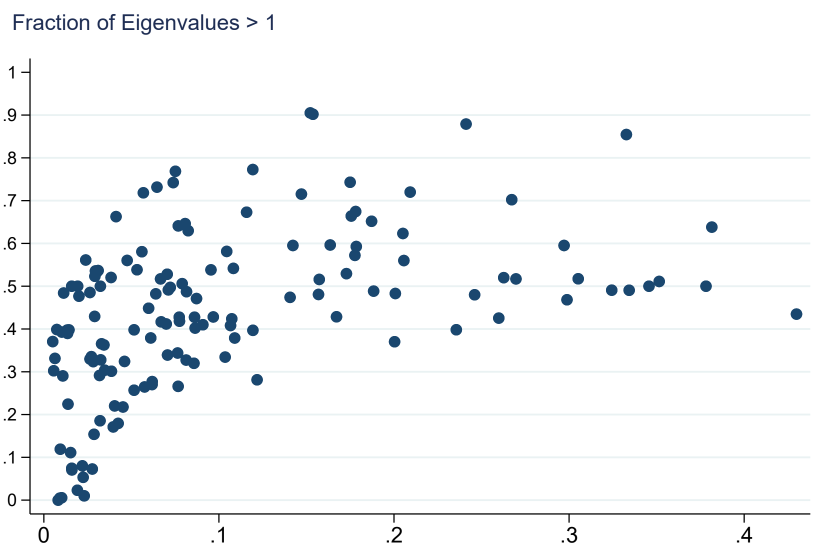
To show the impact of the presence of nonwhite households in a metro area on our estimates of eigenvalues, Figure 2 graphs the share of eigenvalues of $\mathcal{G}$ that are larger than 1 in each MSA on the y-axis against the percentage of black households (top panel) or percentage of hispanic households (bottom panel) in each MSA on the x-axis. As shown in the top panel, on average the share of eigenvalues larger than 1 increases rapidly with the black share of the MSA population until the black share is about 20%, at which point the share of eigenvalues larger than 1 stabilizes at about 50%. Nearly the exact same relationship holds in the bottom panel. When we regress the share of eigenvalues of $\mathcal{G}$ of an MSA greater than 1 on the black and hispanic shares of that MSA, and restrict the sample to MSAs where the black and hispanic shares are each less than 20% (101 out of 141 MSAs), the coefficients on both shares are positive and statistically significant with t-statistics of 13.85 and 15.31, respectively, and the R-squared is 0.82.

The racial composition of neighborhoods at the steady state implied by the current data is unstable because many households want to live in more segregated neighborhoods. This result is not merely a statement about the direction of racial preferences; it is more of a statement about the size of these preferences. To show this, we recompute eigenvalues of $\mathcal{G}$ holding $\delta_j^\tau$ fixed for all tracts $j$ and types $\tau$, but after multiplying all coefficients on race in utility – $a_k^\tau$ for $k = 1, \ldots, 5$ (see equation 9) – by 0.25 for all types. By holding $\delta_j^\tau$ fixed, we preserve the relative desirability of all tracts in the baseline, so any changes to eigenvalues only reflect changes in the strength of preferences for race. The results for all MSAs are shown in column (7) of Appendix table A.1. The bottom line is that with these scaled-down preferences for race, stability for all MSAs vastly improves. Measured at the median MSA, with this rescaling only 3.2% of an MSA's eigenvalues are larger than 1, shown at the end of the table.

Figure 2: Percent of Metro Area Eigenvalues > 1




(a) Black Share of Metro Area




(b) Hispanic Share of Metro Area

# 6 Example Implication for Policy

In this section, we use simulations of the model to predict how neighborhoods will endogenously change in response to a somewhat small policy change that simultaneously affects a relatively large number of tracts. Specifically, for each MSA, we simluate the long-run steady-state predicted response if local governments unexpectedly allow a one-time and immediate 10 percent expansion of all housing developments previously financed using Low Income Housing Tax Credits (LIHTC).[22] The thought behind this analysis was to ask if local governments could implement a relatively small place-based policy in many locations at once without causing a lot of disruption. If the policy was sufficiently small, and implemented in enough locations that already had experience with government policy via LIHTC developments, perhaps incumbent residents would not move in response to a small influx of low-credit-score new residents that may be of a different average racial mix than existing residents.

We implement this counterfactual policy as follows. Denote $\Delta H$ as the total number of new LIHTC units that will be built in the MSA as a consequence of this policy. In the first step, we remove $\Delta H$ housing units (in total) from tracts that are currently housing low-credit-score households in the MSA.[23] Then, in the second step we simulate the model for 5 periods holding $\delta_j$ fixed and $r_j$ fixed. After these 5 periods, we compute births and deaths needed to keep the data (with these $\Delta H$ units removed) in a steady state, before adding the new LIHTC units. Finally, in the third step we add new LIHTC units in proportion to existing LIHTC units until $\Delta H$ units are added.[24] We assume the distribution of types in these new units is the same as the distribution of types from the $\Delta H$ units removed in the first step. With these three steps, we preserve the MSA-wide distribution of types and maintain the MSA-wide aggregate stock of housing, but move $\Delta H$ low-credit score households from tracts without LIHTC units to tracts with LIHTC units. Importantly, the mix of household types moving into the $\Delta H$ new LIHTC units is unlikely the same as the mix of household types in the tracts where those units are located.

Once we have taken the three steps listed above, we compute a new steady state for each MSA. A steady state has the features that (i) the mix of household types in each tract is stable, (ii) the rent in each tract is stable, (iii) the shares of black and hispanic households in each tract is stable, and (iv) expected future rents and black and hispanic shares in each tract are equal to realized shares. When households have strong preferences over the demographic

[22]As we show in Appendix Table A.2, in many MSAs LIHTC developments are located in about 25-35 percent of Census tracts.

[23]The housing units are removed in proportion to the low-credit score population of each tract.

[24]**We use XXXX to identify the location and number of LIHTC units in each metro area.**

composition of their neighborhood, we cannot rule out the possibility that there may be multiple feasible steady states in each MSA. We try to compute a steady state that will be as close to the data as possible, and our algorithm to do this is as follows:

a. Denote the starting total number of households and the rental price in each tract as $\mathcal{H}_j$ and $r_j$, respectively.

b. Do steps 1-5 of the procedure to compute $g(\mathcal{T})$ (see section 5). This generates new simulated black and hispanic shares in each tract, $S_j^{b'}$ and $S_j^{h'}$.

c. Update expected black and hispanic shares in each tract using a weighted average, such that

$$\begin{aligned} \text{new } E\left[S_j^b\right] &= 0.99 * \left(\text{old } E\left[S_j^b\right]\right) + 0.01 * S_j^{b'} \\ \text{new } E\left[S_j^h\right] &= 0.99 * \left(\text{old } E\left[S_j^h\right]\right) + 0.01 * S_j^{h'} \end{aligned}$$

d. Using equation (8) and (9), adjust $\delta_j^\tau$ appropriately given the new values of $E\left[S_j^b\right]$, $E\left[S_j^h\right]$ and $r_j'$, holding exogenous amenities $A_j^\tau$ fixed.

e. Set $r_j = r_j'$ and $\mathcal{H}_j = \mathcal{H}_j'$.

f. Call the sequence of steps b-e as one iteration. Repeat steps b-e *until the distribution of types in each tracts does not change with one additional iteration.*

In the simulations, we tract three statistics for each metro area. The first statistic we compute is the share of tracts that "tip." We define a tract to have tipped if either the black share or the hispanic share changes by 5 percentage points or more in the new steady state relative to the baseline steady state. The other two statistics we compute are black-white and hispanic-white dissimilarity indices. For each metro $m$, we compute these inidices as

$$\begin{aligned} \text{black-white dissimilarity} &= \frac{1}{2}\sum_{j\in m}\left|\frac{b_{j,m}}{B_m} - \frac{w_{j,m}}{W_m}\right| \\ \text{hispanic-white dissimilarity} &= \frac{1}{2}\sum_{j\in m}\left|\frac{h_{j,m}}{H_m} - \frac{w_{j,m}}{W_m}\right| \end{aligned}$$

where $b_{j,m}$, $h_{j,m}$ and $w_{j,m}$ are the numbers of black, hispanic, and white households in tract $j$ of metro $m$ and $B_m$, $H_m$, and $W_m$ are the numbers of black, hispanic, and white households in metro $m$. If there is perfect mixing of races in each tract, then these indices will equal 0; and if there is perfect segregation then the indices will equal 1.

Appendix Table A.2 lists all of the results for each metro area. Column (1) shows the name of the metro and column (2) shows the percentage of tracts with some LIHTC units. Column (3) shows the percentage of tracts that tip in this counterfactual experiment at our baseline estimate of preferences. At the median MSA (shown at the end of the table), more than 60% of tracts tip when racial preferences are at our baseline estimates ("Exp 0"). For comparison, column (4) shows the percentage of tracts that tip in these counterfactual experiments when parameters for racial preferences are set to equal 0.25 of the baseline estimates, ("Exp 1"). In this parameterization, the percentage of tracts that tip falls dramatically: at the Median MSA, less than 6% of tracts tip. This confirms intuition from the eigenvalue analysis that existing neighborhood racial composition is not stable, and the lack of stability arises from strong preferences over the racial composition of neighborhoods.

Columns (5) and (7) show the black-white and hispanic-white dissimilarity indices for our metro areas that are consistent with the steady state implied by our data: measured at the median, the black-white dissimilarity index implied by our data is 31% and the hispanic-white dissimilarity index is 21%. Columns (6) and (9) show the percentage point change in the black-white and hispanic-white dissimilarity indexes for the counterfactual experiment at our baseline estimate of preferences. MSAs become enormously more segregated. At the median MSA, the black-white dissimilarity index increases by nearly 50 percentage points and the hispanic-white dissimilarity index increases by nearly 60 percentage points! Households, on net, want to move to more racially segregated neighborhoods.

Finally, columns (7) and (10) show the percentage point change in the black-white and hispanic-white dissimilarity indexes for the counterfactual experiment when parameters determining racial preferences in the model are set equal to 0.25 of the baseline estimates. Shown in the last rows of the table, for most MSAs the black-white and hispanic-white dissimilarity indexes do not change very much: the median change is -0.3 percentage points in the black-white and 0.1 percentage points in the hispanic-white. Related, when preference parameters for race are set to 0.25 their baseline estimates, in approximately 50 percent of the metro areas both the black-white and the hispanic-white dissimilarity indexes fall, implying that the new steady state has more racial integration than compared to the baseline.

# 7 Conclusion

TBD.

# References

Almagro, M., E. Chyn, and B. Stuart (2022): "The Effects of Urban Renewal Programs on Gentrification and Inequality," Working Paper. 11

Bartik, T. J. (1991): *Who Benefits from State and Local Economic Development Policies?* Upjohn Institute. 1, 12

Baum-Snow, N., and L. Han (2022): "The Microgeography of Housing Supply," Working Paper. 25

Bayer, P., R. McMillan, A. Murphy, and C. Timmins (2015): "A Dynamic Model of Demand for Houses and Neighborhoods," Duke University Working Paper. 5

Bishop, K. C. (2012): "A Dynamic Model of Location Choice and Hedonic Valuation," Working Paper, Washington University in St. Louis. 7

Board of Governors of the Federal Reserve System (2007): "Report to Congress on Credit Scoring and Its Effects on the Availability and Affordability of Credit," . 9

Caetano, G., and V. Maheshri (2021): "A Unified Empirical Framework to Study Segregation," Working Paper. 3, 11

Card, D., A. Mas, and J. Rothstein (2008): "Tipping and the Dynamics of Segregation," *Quarterly Journal of Economics*, 123(1), 177–218. 11

Davis, M. A., J. Gregory, D. A. Hartley, and K. T. K. Tan (2021): "Neighborhood Effects and Housing Vouchers," *Quantitative Economics*, 12(4), 1307–1346. 3, 5, 8, 11

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020): "Bartik Instruments: What, When, Why, and How," *American Economic Review*, 110(8), 2586–2624. 15

Hotz, V. J., and R. A. Miller (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 60(3), 497–529. 7

Kennan, J., and J. R. Walker (2011): "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 79(1), 211–251. 5

# Appendix Table A.1

| Name | Pop (000s) | Tracts | % Black | % Hisp | % Tracts Eigenvalues > 0 Baseline | % Tracts Eigenvalues > 0 $0.25a_k^\tau$ |
|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Akron, OH | 813 | 166 | 0.8% | 10.4% | 33.4% | 3.0% |
| Albany-Schenectady-Troy, NY | 1007 | 214 | 2.3% | 6.2% | 27.7% | 1.2% |
| Albuquerque, NM | 878 | 188 | 43.7% | 3.2% | 50.0% | 4.8% |
| Allentown-Bethlehem-Easton, PA-NJ | 853 | 163 | 6.4% | 2.9% | 42.9% | 0.0% |
| Amarillo, TX | 275 | 62 | 18.5% | 5.6% | 58.1% | 3.2% |
| Anchorage, AK | 367 | 68 | 4.9% | 6.0% | 44.9% | 0.0% |
| Ann Arbor, MI | 377 | 97 | 2.7% | 11.9% | 39.7% | 4.1% |
| Asheville, NC | 487 | 78 | 2.6% | 3.9% | 30.1% | 0.6% |
| Atlantic City-Hammonton, NJ | 315 | 62 | 11.4% | 17.8% | 67.5% | 8.7% |
| Augusta-Richmond County, GA-SC | 611 | 95 | 2.5% | 34.6% | 50.0% | 30.0% |
| Bakersfield, CA | 790 | 136 | 40.4% | 6.5% | 73.2% | 11.8% |
| Barnstable Town, MA | 253 | 50 | 1.3% | 2.3% | 1.0% | 0.0% |
| Baton Rouge, LA | 817 | 142 | 1.8% | 29.9% | 46.8% | 25.0% |
| Beaumont-Port Arthur, TX | 459 | 100 | 7.5% | 20.9% | 72.0% | 14.0% |
| Binghamton, NY | 291 | 65 | 1.7% | 2.9% | 15.4% | 0.0% |
| Birmingham-Hoover, AL | 1234 | 226 | 1.7% | 23.6% | 39.8% | 8.6% |
| Boise City, ID | 539 | 79 | 8.3% | 0.7% | 39.9% | 0.0% |
| 37764 | 799 | 156 | 9.4% | 2.8% | 32.4% | 0.6% |
| Rockingham County-Strafford County, NH | 440 | 79 | 1.3% | 0.8% | 0.0% | 0.0% |
| Boulder, CO | 309 | 62 | 10.9% | 1.1% | 29.0% | 0.0% |
| 14600 | 688 | 143 | 7.2% | 7.1% | 33.9% | 2.1% |
| Bremerton-Silverdale, WA | 257 | 51 | 4.2% | 3.4% | 36.3% | 0.0% |
| Bridgeport-Stamford-Norwalk, CT | 992 | 209 | 11.4% | 10.4% | 58.1% | 6.9% |
| Brownsville-Harlingen, TX | 408 | 86 | 83.3% | 0.6% | 30.2% | 0.6% |
| Canton-Massillon, OH | 467 | 87 | 0.9% | 6.2% | 27.0% | 1.1% |
| Cape Coral-Fort Myers, FL | 574 | 117 | 10.0% | 6.7% | 51.7% | 4.3% |
| Cedar Rapids, IA | 266 | 55 | 1.4% | 2.8% | 7.3% | 0.0% |
| Champaign-Urbana, IL | 253 | 50 | 2.5% | 9.1% | 41.0% | 4.0% |
| Charleston, WV | 357 | 76 | 0.5% | 4.0% | 17.1% | 1.3% |
| Charleston-North Charleston, SC | 748 | 117 | 2.4% | 27.0% | 51.7% | 27.4% |
| Chattanooga, TN-GA | 589 | 98 | 1.3% | 10.7% | 40.8% | 7.7% |
| Gary, IN | 802 | 147 | 9.2% | 14.2% | 59.5% | 4.8% |
| Lake County-Kenosha County, IL-WI | 964 | 181 | 12.3% | 7.2% | 49.7% | 4.1% |
| Clarksville, TN-KY | 324 | 50 | 6.0% | 20.6% | 56.0% | 10.0% |
| Colorado Springs, CO | 635 | 117 | 11.0% | 7.7% | 64.1% | 3.0% |
| Columbia, SC | 839 | 144 | 2.8% | 30.5% | 51.7% | 20.5% |
| Columbus, GA-AL | 396 | 76 | 5.2% | 38.2% | 63.8% | 27.6% |
| Corpus Christi, TX | 500 | 83 | 51.9% | 4.1% | 66.3% | 10.8% |
| Davenport-Moline-Rock Island, IA-IL | 434 | 103 | 5.0% | 5.2% | 39.8% | 1.0% |
| Dayton, OH | 1055 | 208 | 1.1% | 12.2% | 28.1% | 3.1% |
| Deltona-Daytona Beach-Ormond Beach, FL | 537 | 78 | 7.1% | 8.2% | 48.7% | 3.8% |
| Des Moines-West Des Moines, IA | 583 | 107 | 3.5% | 3.5% | 30.4% | 1.4% |
| Duluth, MN-WI | 340 | 90 | 0.8% | 1.0% | 0.6% | 0.0% |
| Durham-Chapel Hill, NC | 521 | 89 | 6.5% | 26.7% | 70.2% | 20.2% |
| El Paso, TX | 880 | 126 | 79.4% | 3.0% | 53.6% | 2.8% |
| Erie, PA | 322 | 72 | 1.8% | 5.2% | 25.7% | 1.4% |
| Eugene, OR | 353 | 78 | 4.6% | 1.4% | 22.4% | 0.0% |
| Evansville, IN-KY | 418 | 85 | 0.9% | 4.5% | 21.8% | 0.6% |
| Fayetteville, NC | 523 | 55 | 9.3% | 33.3% | 85.5% | 40.9% |

| Name | Pop (000s) | Tracts | % Black | % Hisp | % Tracts Eigenvalues > 0 Baseline | $0.25a_k^\tau$ |
|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Fayetteville-Springdale-Rogers, AR-MO | 385 | 68 | 8.2% | 1.4% | 39.0% | 0.0% |
| Flint, MI | 507 | 131 | 2.3% | 20.0% | 37.0% | 6.1% |
| Fort Collins, CO | 294 | 56 | 7.7% | 1.0% | 39.3% | 0.0% |
| Fort Smith, AR-OK | 310 | 52 | 4.2% | 3.3% | 36.5% | 1.0% |
| Fort Wayne, IN | 469 | 104 | 3.2% | 8.7% | 47.1% | 3.4% |
| Fresno, CA | 923 | 158 | 46.3% | 5.7% | 71.8% | 12.3% |
| Grand Rapids-Wyoming, MI | 816 | 159 | 5.5% | 7.7% | 42.8% | 2.8% |
| Green Bay, WI | 336 | 64 | 3.1% | 0.9% | 11.9% | 0.0% |
| Greensboro-High Point, NC | 793 | 142 | 3.9% | 20.5% | 62.3% | 16.5% |
| Greenville-Anderson-Mauldin, SC | 708 | 126 | 2.9% | 15.7% | 51.6% | 8.7% |
| Gulfport-Biloxi-Pascagoula, MS | 322 | 52 | 2.1% | 15.7% | 48.1% | 6.7% |
| Harrisburg-Carlisle, PA | 605 | 111 | 2.6% | 8.6% | 32.0% | 5.0% |
| Hickory-Lenoir-Morganton, NC | 375 | 68 | 3.8% | 7.0% | 41.2% | 0.7% |
| 26180 | 1067 | 212 | 7.0% | 3.2% | 32.8% | 0.0% |
| Huntington-Ashland, WV-KY-OH | 352 | 75 | 0.7% | 2.2% | 5.3% | 0.0% |
| Huntsville, AL | 416 | 87 | 2.0% | 18.8% | 48.9% | 12.6% |
| Jackson, MS | 608 | 115 | 1.0% | 43.0% | 43.5% | 24.8% |
| Jacksonville, FL | 1430 | 201 | 4.0% | 17.8% | 57.2% | 11.2% |
| Kalamazoo-Portage, MI | 367 | 76 | 3.9% | 8.6% | 42.8% | 2.0% |
| Killeen-Temple, TX | 453 | 62 | 15.4% | 24.1% | 87.9% | 21.8% |
| Kingsport-Bristol-Bristol, TN-VA | 356 | 65 | 0.6% | 1.9% | 2.3% | 0.0% |
| Knoxville, TN | 786 | 128 | 1.1% | 4.6% | 32.4% | 2.3% |
| Lafayette, LA | 277 | 50 | 1.6% | 24.6% | 48.0% | 13.0% |
| Lakeland-Winter Haven, FL | 581 | 110 | 9.1% | 11.9% | 77.3% | 6.4% |
| Lancaster, PA | 532 | 94 | 4.4% | 2.6% | 33.0% | 2.1% |
| Lansing-East Lansing, MI | 547 | 117 | 4.0% | 7.1% | 49.1% | 4.3% |
| Lexington-Fayette, KY | 488 | 95 | 2.7% | 10.9% | 37.9% | 2.6% |
| Lincoln, NE | 307 | 62 | 3.2% | 3.2% | 18.5% | 0.0% |
| Little Rock-North Little Rock-Conway, AR | 759 | 147 | 2.1% | 20.1% | 48.3% | 8.5% |
| Lubbock, TX | 303 | 64 | 27.2% | 7.4% | 74.2% | 14.1% |
| Lynchburg, VA | 286 | 56 | 0.9% | 16.7% | 42.9% | 3.6% |
| Macon-Bibb County, GA | 307 | 53 | 1.0% | 33.4% | 49.1% | 20.8% |
| Madison, WI | 572 | 109 | 3.1% | 4.0% | 22.0% | 0.0% |
| Manchester-Nashua, NH | 440 | 81 | 3.0% | 1.5% | 11.1% | 0.0% |
| McAllen-Edinburg-Mission, TX | 653 | 80 | 88.3% | 0.6% | 33.1% | 0.6% |
| Mobile, AL | 507 | 114 | 1.1% | 26.0% | 42.5% | 19.3% |
| Modesto, CA | 495 | 89 | 32.5% | 3.0% | 53.4% | 4.5% |
| Montgomery, AL | 418 | 82 | 1.1% | 37.8% | 50.0% | 28.0% |
| Naples-Immokalee-Marco Island, FL | 291 | 52 | 21.4% | 5.3% | 53.8% | 1.9% |
| New Haven-Milford, CT | 931 | 185 | 8.8% | 11.6% | 67.3% | 8.1% |
| Norwich-New London, CT | 314 | 62 | 5.0% | 6.1% | 37.9% | 2.4% |
| Ogden-Clearfield, UT | 503 | 93 | 8.1% | 1.4% | 39.8% | 0.0% |
| Omaha-Council Bluffs, NE-IA | 856 | 237 | 5.1% | 7.6% | 34.4% | 0.8% |
| Oxnard-Thousand Oaks-Ventura, CA | 891 | 155 | 36.0% | 2.4% | 56.1% | 2.3% |
| Palm Bay-Melbourne-Titusville, FL | 585 | 92 | 4.9% | 8.6% | 40.2% | 1.1% |
| Pensacola-Ferry Pass-Brent, FL | 524 | 77 | 2.8% | 14.1% | 47.4% | 5.2% |
| Peoria, IL | 412 | 94 | 1.4% | 7.7% | 26.6% | 3.2% |
| Wilmington, DE-MD-NJ | 809 | 166 | 4.3% | 16.4% | 59.6% | 10.5% |
| Portland-South Portland, ME | 560 | 108 | 0.9% | 0.9% | 0.5% | 0.0% |
| Port St. Lucie, FL | 389 | 60 | 8.4% | 10.8% | 54.2% | 4.2% |
| 39100 | 705 | 131 | 9.1% | 8.2% | 63.0% | 7.3% |
| Provo-Orem, UT | 423 | 85 | 6.7% | 0.5% | 37.1% | 0.0% |

| Name (1) | Pop (000s) (2) | Tracts (3) | % Black (4) | % Hisp (5) | % Tracts Eigenvalues > 0 Baseline (6) | $0.25a_k^{\tau}$ (7) |
|---|---|---|---|---|---|---|
| Raleigh, NC | 1047 | 128 | 5.3% | 17.6% | 66.4% | 14.5% |
| Reading, PA | 455 | 82 | 6.5% | 3.1% | 53.7% | 3.0% |
| Reno, NV | 419 | 68 | 17.6% | 2.6% | 48.5% | 0.0% |
| Roanoke, VA | 323 | 59 | 1.1% | 10.7% | 42.4% | 5.1% |
| Rockford, IL | 370 | 82 | 7.5% | 8.1% | 64.6% | 4.9% |
| Saginaw, MI | 244 | 56 | 6.3% | 18.7% | 65.2% | 6.3% |
| Salem, OR | 405 | 63 | 16.0% | 1.1% | 48.4% | 0.0% |
| Salinas, CA | 519 | 83 | 53.2% | 4.8% | 56.0% | 6.0% |
| Salt Lake City, UT | 1123 | 205 | 12.3% | 1.3% | 39.8% | 0.2% |
| 42060 | 445 | 86 | 36.3% | 2.9% | 52.3% | 4.7% |
| Santa Cruz-Watsonville, CA | 289 | 52 | 29.5% | 1.6% | 50.0% | 0.0% |
| Santa Rosa, CA | 507 | 86 | 18.1% | 2.0% | 47.7% | 0.0% |
| Savannah, GA | 441 | 75 | 2.2% | 26.3% | 52.0% | 24.0% |
| Scranton–Wilkes-Barre–Hazleton, PA | 668 | 168 | 1.1% | 1.6% | 7.4% | 0.0% |
| Tacoma-Lakewood, WA | 806 | 157 | 6.0% | 9.5% | 53.8% | 0.0% |
| Shreveport-Bossier City, LA | 471 | 90 | 1.8% | 35.1% | 51.1% | 26.1% |
| South Bend-Mishawaka, IN-MI | 402 | 84 | 3.5% | 7.9% | 50.6% | 7.1% |
| Spartanburg, SC | 326 | 51 | 2.5% | 17.3% | 52.9% | 4.9% |
| Spokane-Spokane Valley, WA | 473 | 106 | 2.7% | 2.2% | 8.0% | 0.0% |
| Springfield, IL | 229 | 55 | 0.8% | 7.7% | 41.8% | 2.7% |
| Springfield, MA | 779 | 140 | 9.8% | 6.4% | 48.2% | 5.4% |
| Springfield, MO | 418 | 85 | 1.5% | 1.6% | 7.1% | 0.0% |
| Stockton-Lodi, CA | 659 | 121 | 30.1% | 7.5% | 76.9% | 14.0% |
| Syracuse, NY | 774 | 189 | 1.8% | 5.8% | 26.5% | 1.3% |
| Tallahassee, FL | 385 | 63 | 3.7% | 29.7% | 59.5% | 23.0% |
| Toledo, OH | 777 | 174 | 4.2% | 9.7% | 42.8% | 4.0% |
| Topeka, KS | 254 | 54 | 5.0% | 6.7% | 41.7% | 0.9% |
| Trenton, NJ | 402 | 72 | 8.4% | 17.5% | 74.3% | 13.2% |
| Tucson, AZ | 1026 | 196 | 30.1% | 3.8% | 52.0% | 2.0% |
| Tuscaloosa, AL | 232 | 54 | 1.2% | 32.4% | 49.1% | 22.2% |
| Utica-Rome, NY | 369 | 92 | 2.4% | 4.2% | 17.9% | 1.1% |
| Vallejo-Fairfield, CA | 478 | 79 | 17.3% | 15.2% | 90.5% | 18.4% |
| Visalia-Porterville, CA | 419 | 76 | 52.3% | 1.9% | 50.0% | 6.6% |
| Waco, TX | 238 | 51 | 18.3% | 15.4% | 90.2% | 14.7% |
| 13644 | 1263 | 209 | 9.4% | 14.7% | 71.5% | 15.8% |
| Wichita, KS | 686 | 143 | 6.6% | 7.0% | 52.8% | 2.4% |
| Winston-Salem, NC | 483 | 97 | 5.6% | 17.8% | 59.3% | 11.3% |
| Worcester, MA-CT | 882 | 163 | 6.2% | 3.2% | 29.1% | 0.3% |
| York-Hanover, PA | 451 | 82 | 2.1% | 2.7% | 33.5% | 2.4% |
| Youngstown-Warren-Boardman, OH-PA | 748 | 168 | 1.4% | 8.1% | 32.7% | 1.5% |
| 25th Percentile | 367 | 68 | 2.1% | 3.0% | 33.0% | 0.6% |
| Median | 483 | 89 | 4.4% | 7.1% | 46.8% | 3.2% |
| 75th Percentile | 748 | 131 | 9.2% | 15.4% | 53.8% | 8.7% |

# Appendix Table A.2

| Name | LIHTC % | Tracts Tipping | | Black-White Dissim | | | Hisp-White Dissim | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Exp 0 | Exp 1 | Base | Δ 0 | Δ 1 | Base | Δ 0 | Δ 1 |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Akron, OH | 28.4% | 56.1% | 8.9% | 43.5% | 47.8 pp | 0.1 pp | 16.5% | 71.4 pp | 2.2 pp |
| Albany-Schenectady-Troy, NY | 22.3% | 24.1% | 2.2% | 38.3% | 49.2 pp | -2.5 pp | 21.7% | 62.8 pp | -0.4 pp |
| Albuquerque, NM | 23.8% | 93.3% | 29.7% | 15.3% | 61.7 pp | -0.7 pp | 26.7% | 59.4 pp | 0.9 pp |
| Allentown-Bethlehem-Easton, PA-NJ | 27.6% | 33.9% | 2.4% | 29.9% | 40.0 pp | -0.8 pp | 36.0% | 47.0 pp | -1.8 pp |
| Amarillo, TX | 24.5% | 67.4% | 7.3% | 32.6% | 46.2 pp | -5.2 pp | 27.5% | 58.1 pp | -2.5 pp |
| Anchorage, AK | 34.2% | 48.2% | 0.0% | 21.2% | 53.7 pp | -0.2 pp | 14.4% | 59.9 pp | 0.0 pp |
| Ann Arbor, MI | 23.7% | 52.9% | 14.1% | 36.2% | 49.9 pp | 1.7 pp | 15.2% | 62.9 pp | 5.2 pp |
| Asheville, NC | 32.7% | 10.8% | 1.2% | 26.5% | 41.2 pp | -1.8 pp | 16.0% | 45.5 pp | 0.0 pp |
| Atlantic City-Hammonton, NJ | 4.6% | 92.0% | 4.6% | 42.0% | 50.4 pp | -3.6 pp | 29.7% | 62.3 pp | -2.5 pp |
| Augusta-Richmond County, GA-SC | 23.4% | 97.8% | 46.0% | 31.8% | 58.2 pp | 2.0 pp | 16.0% | 73.8 pp | 2.5 pp |
| Bakersfield, CA | 36.9% | 92.6% | 51.4% | 29.3% | 61.4 pp | -2.3 pp | 37.7% | 51.2 pp | -1.9 pp |
| Barnstable Town, MA | 20.9% | 0.0% | 0.0% | 18.0% | -0.6 pp | -0.2 pp | 15.4% | 0.1 pp | -0.1 pp |
| Baton Rouge, LA | 30.3% | 83.9% | 43.7% | 43.9% | 46.4 pp | 4.4 pp | 17.3% | 69.7 pp | 4.9 pp |
| Beaumont-Port Arthur, TX | 23.3% | 82.3% | 15.8% | 53.8% | 35.8 pp | -0.4 pp | 28.3% | 59.9 pp | 3.5 pp |
| Binghamton, NY | 31.4% | 0.0% | 0.0% | 26.2% | -0.9 pp | -0.2 pp | 19.4% | -0.7 pp | -0.3 pp |
| Birmingham-Hoover, AL | 31.1% | 81.0% | 25.1% | 52.0% | 37.5 pp | 2.0 pp | 21.5% | 68.1 pp | 6.8 pp |
| Boise City, ID | 43.6% | 46.5% | 1.5% | 9.4% | 2.6 pp | 0.3 pp | 20.0% | 52.2 pp | -1.4 pp |
| 37764 | 24.2% | 28.8% | 3.3% | 32.1% | 36.3 pp | -1.3 pp | 50.3% | 29.0 pp | -1.6 pp |
| Rockingham County-Strafford County, NH | 36.9% | 0.0% | 0.0% | 12.4% | -0.3 pp | -0.1 pp | 15.5% | -0.6 pp | -0.1 pp |
| Boulder, CO | 35.5% | 36.5% | 2.1% | 10.7% | 5.1 pp | -0.2 pp | 22.5% | 42.7 pp | -1.7 pp |
| 14600 | 13.8% | 23.0% | 4.5% | 33.4% | 49.9 pp | -2.5 pp | 23.1% | 57.5 pp | -0.7 pp |
| Bremerton-Silverdale, WA | 38.4% | 4.9% | 0.0% | 20.0% | 10.8 pp | 0.1 pp | 10.3% | 13.2 pp | 0.3 pp |
| Bridgeport-Stamford-Norwalk, CT | 17.1% | 88.1% | 5.4% | 48.2% | 38.5 pp | -0.9 pp | 41.0% | 48.4 pp | -0.6 pp |
| Brownsville-Harlingen, TX | 23.9% | 89.2% | 46.5% | 19.9% | 51.8 pp | 5.0 pp | 31.4% | 47.4 pp | 11.4 pp |
| Canton-Massillon, OH | 21.1% | 28.1% | 1.5% | 36.3% | 37.4 pp | -0.5 pp | 13.1% | 8.2 pp | 0.1 pp |
| Cape Coral-Fort Myers, FL | 12.4% | 58.1% | 1.9% | 35.5% | 51.6 pp | -5.5 pp | 23.9% | 63.6 pp | 1.9 pp |
| Cedar Rapids, IA | 33.3% | 1.4% | 0.0% | 24.8% | 2.2 pp | -0.2 pp | 13.2% | 4.5 pp | 0.6 pp |
| Champaign-Urbana, IL | 25.6% | 50.3% | 7.7% | 34.2% | 56.4 pp | 2.2 pp | 22.7% | 64.2 pp | 3.6 pp |
| Charleston, WV | 33.1% | 12.6% | 0.0% | 35.8% | 28.6 pp | -1.0 pp | 18.1% | 10.4 pp | 0.1 pp |
| Charleston-North Charleston, SC | 34.8% | 94.8% | 50.0% | 27.0% | 62.0 pp | 8.3 pp | 15.8% | 71.8 pp | 8.4 pp |
| Chattanooga, TN-GA | 30.2% | 60.3% | 11.9% | 49.1% | 41.5 pp | -5.8 pp | 16.9% | 64.1 pp | 3.0 pp |
| Gary, IN | 14.9% | 86.3% | 14.8% | 60.9% | 31.6 pp | 0.8 pp | 32.0% | 58.6 pp | 1.7 pp |
| Lake County-Kenosha County, IL-WI | 20.1% | 86.6% | 4.3% | 45.3% | 45.5 pp | -2.7 pp | 39.5% | 52.1 pp | -3.1 pp |
| Clarksville, TN-KY | 25.2% | 89.0% | 10.8% | 26.8% | 62.6 pp | 0.8 pp | 26.1% | 63.6 pp | 0.9 pp |
| Colorado Springs, CO | 13.6% | 59.8% | 2.2% | 23.2% | 51.3 pp | -0.5 pp | 16.3% | 45.5 pp | -0.5 pp |
| Columbia, SC | 34.0% | 94.8% | 49.7% | 35.5% | 56.0 pp | 2.8 pp | 20.1% | 70.7 pp | 6.0 pp |
| Columbus, GA-AL | 16.0% | 100.0% | 15.8% | 41.5% | 50.2 pp | 1.1 pp | 24.1% | 61.0 pp | 3.2 pp |
| Corpus Christi, TX | 23.7% | 97.5% | 54.6% | 21.9% | 64.0 pp | -2.5 pp | 33.1% | 53.9 pp | -1.7 pp |
| Davenport-Moline-Rock Island, IA-IL | 27.2% | 24.7% | 1.8% | 30.3% | 46.3 pp | -0.9 pp | 23.9% | 54.7 pp | -0.7 pp |
| Dayton, OH | 41.6% | 79.9% | 21.8% | 53.1% | 37.6 pp | -0.2 pp | 17.0% | 70.9 pp | 5.6 pp |
| Deltona-Daytona Beach-Ormond Beach, FL | 21.4% | 57.6% | 3.2% | 32.8% | 52.6 pp | -4.6 pp | 28.6% | 56.5 pp | 1.5 pp |
| Des Moines-West Des Moines, IA | 50.5% | 20.0% | 2.0% | 28.8% | 57.3 pp | -2.9 pp | 25.6% | 57.6 pp | -2.1 pp |
| Duluth, MN-WI | 27.0% | 0.0% | 0.0% | 21.7% | -0.2 pp | -0.1 pp | 14.7% | 0.0 pp | 0.0 pp |
| Durham-Chapel Hill, NC | 37.7% | 100.0% | 43.7% | 33.6% | 59.5 pp | 6.8 pp | 29.7% | 36.5 pp | 10.5 pp |
| El Paso, TX | 45.1% | 95.3% | 34.4% | 21.0% | 66.9 pp | 3.1 pp | 30.5% | 51.2 pp | 1.6 pp |
| Erie, PA | 23.0% | 21.9% | 1.0% | 41.7% | 34.9 pp | -0.8 pp | 27.2% | 38.6 pp | 0.0 pp |
| Eugene, OR | 40.5% | 0.0% | 0.0% | 16.4% | -0.5 pp | -0.2 pp | 10.4% | -0.2 pp | -0.4 pp |
| Evansville, IN-KY | 26.1% | 27.9% | 0.6% | 30.6% | 52.7 pp | 0.0 pp | 14.9% | 16.7 pp | 1.0 pp |
| Fayetteville, NC | 28.5% | 100.0% | 47.9% | 20.1% | 72.2 pp | 7.2 pp | 15.3% | 53.4 pp | 5.9 pp |

| Name | LIHTC % | Tracts Tipping | | Black-White Dissim | | | Hisp-White Dissim | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Exp 0 | Exp 1 | Base | Δ 0 | Δ 1 | Base | Δ 0 | Δ 1 |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Fayetteville-Springdale-Rogers, AR-MO | 45.3% | 53.9% | 0.0% | 17.8% | 24.4 pp | -0.5 pp | 24.5% | 55.0 pp | -1.8 pp |
| Flint, MI | 26.2% | 76.1% | 27.2% | 56.6% | 32.1 pp | 2.6 pp | 16.6% | 62.1 pp | 10.7 pp |
| Fort Collins, CO | 43.1% | 7.2% | 0.0% | 11.7% | 4.3 pp | -0.4 pp | 8.4% | 25.9 pp | -0.5 pp |
| Fort Smith, AR-OK | 57.3% | 15.8% | 0.0% | 30.5% | 3.2 pp | -1.6 pp | 30.1% | 3.2 pp | -1.8 pp |
| Fort Wayne, IN | 28.9% | 44.2% | 11.0% | 42.3% | 46.5 pp | -3.3 pp | 24.9% | 63.4 pp | 0.2 pp |
| Fresno, CA | 43.8% | 94.5% | 61.1% | 29.7% | 61.1 pp | -3.8 pp | 30.7% | 58.6 pp | 2.9 pp |
| Grand Rapids-Wyoming, MI | 28.8% | 35.7% | 2.1% | 42.5% | 45.7 pp | -0.9 pp | 32.0% | 56.9 pp | -0.5 pp |
| Green Bay, WI | 38.9% | 0.0% | 0.0% | 23.3% | -0.3 pp | -0.1 pp | 27.7% | -1.0 pp | -0.3 pp |
| Greensboro-High Point, NC | 33.6% | 97.1% | 27.8% | 37.8% | 54.2 pp | 3.9 pp | 23.6% | 56.4 pp | 6.6 pp |
| Greenville-Anderson-Mauldin, SC | 38.2% | 70.5% | 16.3% | 29.6% | 57.9 pp | 0.4 pp | 17.9% | 69.1 pp | 2.9 pp |
| Gulfport-Biloxi-Pascagoula, MS | 59.3% | 68.7% | 33.8% | 27.1% | 58.6 pp | -0.1 pp | 16.6% | 66.7 pp | 9.9 pp |
| Harrisburg-Carlisle, PA | 33.7% | 42.5% | 3.9% | 49.0% | 39.5 pp | -0.6 pp | 28.8% | 47.8 pp | 1.9 pp |
| Hickory-Lenoir-Morganton, NC | 31.5% | 28.2% | 1.3% | 23.4% | 53.0 pp | -1.1 pp | 19.1% | 57.2 pp | -0.1 pp |
| 26180 | 16.3% | 0.7% | 0.0% | 24.9% | -1.0 pp | -0.2 pp | 14.1% | -0.9 pp | -0.2 pp |
| Huntington-Ashland, WV-KY-OH | 29.2% | 2.2% | 0.0% | 31.6% | 0.4 pp | -1.0 pp | 16.7% | 1.8 pp | -0.1 pp |
| Huntsville, AL | 29.1% | 80.3% | 31.6% | 35.3% | 55.5 pp | 5.0 pp | 17.3% | 69.4 pp | 7.7 pp |
| Jackson, MS | 47.5% | 91.8% | 46.2% | 42.8% | 47.2 pp | 5.8 pp | 17.3% | 75.7 pp | 18.3 pp |
| Jacksonville, FL | 35.2% | 85.7% | 29.8% | 37.0% | 52.6 pp | 3.9 pp | 14.2% | 76.1 pp | 7.3 pp |
| Kalamazoo-Portage, MI | 44.2% | 48.5% | 4.1% | 31.1% | 57.0 pp | -2.7 pp | 23.4% | 64.6 pp | 1.1 pp |
| Killeen-Temple, TX | 17.5% | 97.8% | 6.7% | 30.2% | 58.1 pp | -1.8 pp | 14.2% | 73.6 pp | -0.3 pp |
| Kingsport-Bristol-Bristol, TN-VA | 28.4% | 2.0% | 0.0% | 22.9% | 0.6 pp | -0.4 pp | 13.0% | 1.7 pp | 0.3 pp |
| Knoxville, TN | 25.4% | 23.3% | 1.7% | 32.0% | 56.1 pp | -1.2 pp | 13.5% | 70.3 pp | 0.4 pp |
| Lafayette, LA | 25.1% | 92.3% | 28.9% | 33.3% | 55.5 pp | 3.3 pp | 12.7% | 59.9 pp | 2.8 pp |
| Lakeland-Winter Haven, FL | 14.9% | 78.1% | 7.7% | 26.6% | 58.9 pp | 0.1 pp | 17.8% | 67.8 pp | 3.7 pp |
| Lancaster, PA | 25.1% | 17.7% | 4.3% | 31.5% | 49.7 pp | -2.2 pp | 36.5% | 50.7 pp | -3.7 pp |
| Lansing-East Lansing, MI | 36.3% | 39.4% | 2.1% | 37.9% | 39.5 pp | -1.9 pp | 23.9% | 43.6 pp | -0.3 pp |
| Lexington-Fayette, KY | 19.5% | 28.3% | 2.8% | 31.0% | 50.4 pp | -2.9 pp | 15.6% | 67.1 pp | -0.5 pp |
| Lincoln, NE | 33.3% | 2.4% | 0.0% | 20.2% | 4.5 pp | 0.1 pp | 16.6% | 7.2 pp | 0.3 pp |
| Little Rock-North Little Rock-Conway, AR | 35.5% | 79.8% | 40.2% | 41.0% | 49.3 pp | 5.8 pp | 17.0% | 68.2 pp | 12.5 pp |
| Lubbock, TX | 22.4% | 91.6% | 36.3% | 37.4% | 49.9 pp | -7.6 pp | 29.7% | 60.3 pp | 1.1 pp |
| Lynchburg, VA | 24.5% | 84.9% | 11.1% | 27.8% | 60.4 pp | 0.6 pp | 14.0% | 72.0 pp | 4.4 pp |
| Macon-Bibb County, GA | 22.3% | 96.2% | 22.4% | 39.7% | 48.6 pp | 3.0 pp | 14.9% | 28.7 pp | 8.5 pp |
| Madison, WI | 63.2% | 16.2% | 0.0% | 27.0% | 38.5 pp | -0.4 pp | 18.9% | 62.8 pp | 0.3 pp |
| Manchester-Nashua, NH | 33.2% | 3.4% | 0.0% | 16.7% | 8.4 pp | -0.3 pp | 22.0% | 24.0 pp | -0.7 pp |
| McAllen-Edinburg-Mission, TX | 41.4% | 97.6% | 39.3% | 14.2% | 57.0 pp | 5.9 pp | 20.6% | 55.1 pp | 2.6 pp |
| Mobile, AL | 27.7% | 83.9% | 51.1% | 45.3% | 42.7 pp | 6.3 pp | 15.1% | 71.4 pp | 15.9 pp |
| Modesto, CA | 23.5% | 97.8% | 9.0% | 18.2% | 42.1 pp | 0.5 pp | 21.1% | 62.0 pp | -0.9 pp |
| Montgomery, AL | 42.6% | 90.8% | 49.4% | 39.0% | 52.5 pp | 4.8 pp | 16.8% | 75.2 pp | 11.9 pp |
| Naples-Immokalee-Marco Island, FL | 32.2% | 52.9% | 16.5% | 29.5% | 51.5 pp | -2.6 pp | 35.5% | 45.0 pp | -5.5 pp |
| New Haven-Milford, CT | 20.5% | 84.8% | 7.5% | 46.9% | 43.4 pp | -3.2 pp | 38.4% | 52.0 pp | -1.5 pp |
| Norwich-New London, CT | 20.3% | 19.6% | 0.9% | 32.1% | 40.3 pp | -1.7 pp | 27.6% | 49.9 pp | -1.4 pp |
| Ogden-Clearfield, UT | 21.9% | 48.0% | 2.0% | 20.0% | 2.6 pp | -0.3 pp | 20.6% | 46.6 pp | -1.2 pp |
| Omaha-Council Bluffs, NE-IA | 33.1% | 42.4% | 0.0% | 42.6% | 46.0 pp | 0.0 pp | 28.3% | 58.4 pp | 0.0 pp |
| Oxnard-Thousand Oaks-Ventura, CA | 24.4% | 79.9% | 29.2% | 23.2% | 51.2 pp | -2.5 pp | 41.0% | 44.9 pp | -6.1 pp |
| Palm Bay-Melbourne-Titusville, FL | 13.2% | 32.7% | 1.0% | 27.1% | 42.3 pp | -0.1 pp | 13.4% | 52.9 pp | 1.0 pp |
| Pensacola-Ferry Pass-Brent, FL | 22.3% | 65.1% | 9.9% | 31.1% | 51.8 pp | 0.3 pp | 12.7% | 61.6 pp | 0.8 pp |
| Peoria, IL | 20.3% | 25.1% | 3.0% | 48.2% | 39.1 pp | -0.3 pp | 19.9% | 60.9 pp | 1.0 pp |
| Wilmington, DE-MD-NJ | 20.8% | 93.2% | 24.3% | 36.7% | 54.7 pp | 2.1 pp | 26.9% | 64.4 pp | 4.3 pp |
| Portland-South Portland, ME | 42.8% | 0.0% | 0.0% | 18.6% | -0.1 pp | -0.1 pp | 12.0% | 0.4 pp | 0.0 pp |
| Port St. Lucie, FL | 20.9% | 86.8% | 14.2% | 31.4% | 55.2 pp | 4.3 pp | 20.7% | 29.3 pp | 9.7 pp |
| 39100 | 34.6% | 86.8% | 4.5% | 31.7% | 55.9 pp | -3.5 pp | 23.2% | 64.7 pp | -1.4 pp |
| Provo-Orem, UT | 16.3% | 67.8% | 0.0% | 13.5% | 15.5 pp | -0.2 pp | 13.9% | 70.1 pp | -0.2 pp |

| | | Tracts Tipping | | Black-White Dissim | | | Hisp-White Dissim | | |
|---|---|---|---|---|---|---|---|---|---|
| Name | LIHTC % | Exp 0 | Exp 1 | Base | Δ 0 | Δ 1 | Base | Δ 0 | Δ 1 |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Raleigh, NC | 55.0% | 100.0% | 33.1% | 26.5% | 67.6 pp | 5.2 pp | 15.5% | 59.6 pp | 9.4 pp |
| Reading, PA | 21.6% | 27.5% | 6.1% | 34.1% | 53.2 pp | -0.5 pp | 47.8% | 41.8 pp | -1.9 pp |
| Reno, NV | 38.7% | 95.7% | 7.7% | 13.9% | 18.2 pp | -0.9 pp | 20.2% | 62.2 pp | -1.3 pp |
| Roanoke, VA | 26.3% | 51.4% | 5.5% | 41.4% | 48.2 pp | -2.5 pp | 15.9% | 59.9 pp | 2.4 pp |
| Rockford, IL | 19.1% | 48.5% | 2.2% | 37.7% | 50.5 pp | 0.1 pp | 25.1% | 63.9 pp | 0.2 pp |
| Saginaw, MI | 31.8% | 75.6% | 14.3% | 55.0% | 36.3 pp | -2.4 pp | 32.8% | 57.4 pp | 1.6 pp |
| Salem, OR | 36.5% | 96.4% | 2.9% | 12.9% | 8.6 pp | -0.7 pp | 25.4% | 57.2 pp | -3.1 pp |
| Salinas, CA | 36.2% | 83.5% | 62.6% | 29.9% | 57.2 pp | -0.1 pp | 46.3% | 42.2 pp | -4.6 pp |
| Salt Lake City, UT | 28.2% | 76.1% | 3.3% | 16.2% | 12.5 pp | -0.8 pp | 25.1% | 58.8 pp | -1.0 pp |
| 42060 | 32.0% | 84.5% | 22.4% | 22.6% | 51.2 pp | -1.5 pp | 29.7% | 53.5 pp | -2.0 pp |
| Santa Cruz-Watsonville, CA | 37.6% | 71.8% | 18.2% | 12.0% | 56.2 pp | -0.7 pp | 44.8% | 41.4 pp | -8.7 pp |
| Santa Rosa, CA | 46.7% | 93.6% | 7.9% | 15.2% | 55.7 pp | -1.4 pp | 18.1% | 58.6 pp | -3.2 pp |
| Savannah, GA | 18.0% | 95.5% | 26.5% | 37.9% | 54.9 pp | -0.8 pp | 18.1% | 27.2 pp | 4.5 pp |
| Scranton–Wilkes-Barre–Hazleton, PA | 9.9% | 0.0% | 0.0% | 28.4% | -0.6 pp | -0.1 pp | 29.4% | -0.6 pp | -0.2 pp |
| Tacoma-Lakewood, WA | 23.7% | 59.1% | 0.0% | 26.4% | 55.5 pp | -0.2 pp | 14.7% | 68.1 pp | 0.1 pp |
| Shreveport-Bossier City, LA | 45.1% | 91.7% | 51.8% | 42.0% | 48.3 pp | 6.2 pp | 15.0% | 57.7 pp | 9.8 pp |
| South Bend-Mishawaka, IN-MI | 20.8% | 54.9% | 2.8% | 37.9% | 52.0 pp | -0.8 pp | 28.4% | 59.2 pp | -0.6 pp |
| Spartanburg, SC | 34.7% | 94.9% | 16.0% | 25.1% | 64.8 pp | 2.6 pp | 17.0% | 72.3 pp | 1.6 pp |
| Spokane-Spokane Valley, WA | 33.3% | 0.0% | 0.0% | 13.8% | -0.3 pp | -0.1 pp | 8.9% | -0.6 pp | -0.1 pp |
| Springfield, IL | 24.5% | 26.6% | 2.2% | 36.7% | 47.7 pp | -1.9 pp | 15.5% | 58.7 pp | -0.1 pp |
| Springfield, MA | 29.0% | 53.3% | 9.2% | 48.5% | 38.6 pp | -6.0 pp | 46.5% | 45.1 pp | -4.6 pp |
| Springfield, MO | 52.1% | 0.0% | 0.0% | 18.4% | -0.5 pp | -0.2 pp | 11.2% | -0.2 pp | -0.1 pp |
| Stockton-Lodi, CA | 18.7% | 95.2% | 14.8% | 26.5% | 59.5 pp | -1.6 pp | 18.9% | 67.8 pp | -0.1 pp |
| Syracuse, NY | 20.9% | 26.6% | 2.7% | 46.3% | 44.8 pp | -1.2 pp | 25.1% | 61.0 pp | 0.2 pp |
| Tallahassee, FL | 31.1% | 99.7% | 40.8% | 31.4% | 59.5 pp | 2.2 pp | 14.9% | 74.6 pp | 6.4 pp |
| Toledo, OH | 26.2% | 39.9% | 6.2% | 49.0% | 39.2 pp | 0.2 pp | 20.9% | 65.2 pp | 1.6 pp |
| Topeka, KS | 53.2% | 35.0% | 1.4% | 31.5% | 45.0 pp | -2.1 pp | 25.4% | 48.3 pp | -1.4 pp |
| Trenton, NJ | 10.5% | 98.5% | 5.3% | 48.1% | 44.6 pp | 0.1 pp | 32.8% | 49.1 pp | 1.0 pp |
| Tucson, AZ | 16.4% | 80.1% | 10.8% | 19.2% | 50.3 pp | -0.5 pp | 34.0% | 49.6 pp | -2.6 pp |
| Tuscaloosa, AL | 42.0% | 97.3% | 58.8% | 38.6% | 53.7 pp | 4.1 pp | 16.4% | 76.7 pp | 9.3 pp |
| Utica-Rome, NY | 22.5% | 12.9% | 1.7% | 39.6% | 40.4 pp | -2.3 pp | 31.7% | 39.1 pp | -1.1 pp |
| Vallejo-Fairfield, CA | 31.0% | 96.3% | 1.4% | 22.8% | 60.6 pp | -1.7 pp | 16.1% | 68.2 pp | 1.1 pp |
| Visalia-Porterville, CA | 45.8% | 99.8% | 86.2% | 19.3% | 56.0 pp | 3.0 pp | 24.9% | 63.0 pp | -1.7 pp |
| Waco, TX | 16.5% | 95.7% | 13.8% | 36.8% | 51.1 pp | -8.7 pp | 30.0% | 54.7 pp | 2.1 pp |
| 13644 | 24.3% | 97.9% | 30.4% | 29.0% | 61.5 pp | 2.7 pp | 29.4% | 61.8 pp | 4.1 pp |
| Wichita, KS | 39.7% | 70.1% | 3.3% | 38.0% | 52.0 pp | -2.8 pp | 23.3% | 66.7 pp | -0.8 pp |
| Winston-Salem, NC | 30.3% | 96.6% | 20.4% | 41.0% | 49.5 pp | 0.3 pp | 27.2% | 13.8 pp | 3.3 pp |
| Worcester, MA-CT | 23.0% | 34.7% | 2.6% | 29.9% | 1.3 pp | -1.0 pp | 34.3% | 48.2 pp | -1.4 pp |
| York-Hanover, PA | 27.6% | 17.8% | 1.3% | 28.0% | 58.5 pp | -1.9 pp | 28.4% | 59.5 pp | -1.2 pp |
| Youngstown-Warren-Boardman, OH-PA | 25.9% | 34.5% | 5.6% | 48.5% | 41.3 pp | -1.3 pp | 29.0% | 55.6 pp | 0.3 pp |
| 25th Percentile | 22.3% | 27.5% | 1.5% | 23.2% | 37.4 pp | -1.6 pp | 16.0% | 44.9 pp | -0.8 pp |
| Median | 28.4% | 60.3% | 5.6% | 31.1% | 48.6 pp | -0.3 pp | 21.1% | 57.4 pp | 0.1 pp |
| 75th Percentile | 35.5% | 91.7% | 22.9% | 38.3% | 55.5 pp | 0.5 pp | 28.4% | 63.6 pp | 2.9 pp |