The Long-Run Effects of Low-Income Housing on Neighborhood Composition*

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

July 3, 2019

Abstract

We develop a new model of the demand for neighborhoods and use the model to forecast the long-run impact of new low-income housing units on neighborhood demographic composition and housing rents. We estimate the utility that each of a large number of observable "types" of households derive from neighborhoods (Census tracts) in MSAs throughout the U.S. using detailed panel data on the location choices of 5% of the U.S. population. We then estimate each type's preferences over the share of low-income and black residents of the neighborhood, exploiting a new instrumental variables approach that combines the implications of our model with two discontinuities in the formula used by the the department of Housing and Urban Development (HUD) for determining eligibility for federal low-income housing development credits. With knowledge of each type's preferences for neighborhoods and demographics, we simulate the impact of newly built low-income housing units on the long-run level of rent and the share of black and low-income residents in the tracts receiving the units. Finally, we combine the new Opportunity Atlas data set of Chetty, Friedman, Hendren, Jones, and Porter (2018) with simulations of our model to study the degree to which newly built low-income housing units impacts the adult earnings of children.

JEL Classification Numbers: Insert classifications here
Keywords: Housing Vouchers, 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

We study how targeted low-income housing development projects change the long-run racial and economic composition of neighborhoods when households have preferences over the race and income of their neighbors. When first built, new low-income housing development adds low-income neighbors which, for many neighborhoods, increases both racial and economic diversity of the neighborhood. However, some households with sufficiently strong preferences move in response to the increased diversity. Neighborhood composition changes as households with different preferences over the race and income of their neighbors move in and out. Rents adjust to clear markets and this generates additional migration of households sensitive to rental prices. When all is said and done, the long-run consequences of neighborhoold composition resulting from new low-income housing units depends on how people move and how rents adjust.

To study these issues, we construct and estimate a large-scale model of housing demand for all neighborhoods in a metropolitan area. The model contains all the ingredients we believe are key to understanding how neighborhoods change. Households have preferences for neighborhoods directly, but also have preference for the racial and socio-economic characteristics of the residents of their neighborhood. Additionally, households care about rent. Households differ with respect to their preferences for neighborhoods, neighbors and rents. Given these preferences and the racial and socio-economic makeup of each neighborhood in their metro area of residence, they optimally choose where to live.

When low-income housing is added to a neighborhood, the desirability of that neighborhood relative to other neighborhoods in the metro area may change due to changes in the demographic characteristics brought on by the new low-income housing units. Some people might move in and others move out, but migration occurs slowly as it is costly to move. Over time, the racial and socio-economic composition of each neighborhood in the metro area may change as people move. Additionally, rents may also adjust to ensure the demand for space in each neighborhood is equal to the supply. Simulations of the model allow us to directly predict how households move and how neighborhoods demographics and rents change in response to the addition of new low-income housing units in any given neighborhood or set of neighborhoods in a metro area.

Our results depend on our estimates of (1) how individual households directly value neighborhoods and how this value changes with rent as well as the demographic and socio-economic composotion of that neighborhood composition as well as (2) how those preferences vary across households inside each metro area. Conceptually, we estimate these preferences and model parameters in two-steps. In the first step, we estimate preferences for all locations assuming that people assume the level of rent and socio-economic and demographic makeup of each location is fixed at a baseline level. These estimates are derived from annual data on the location choices from 1999 to present of 5% of the U.S. population from the NYFRB/Equifax Consumer Credit Panel. We impose that a neighborhood in this model corresponds to a Census tract in these data. We estimate baseline preferences for living in every Census tract in every U.S. Metropolitan Statistical Area by maximum likelihood, and we allow these baseline preferences to vary across 315 fixed "types" of households. Households are sorted into types based on their observable characteristics (credit score, for example) the first time they are observed in the sample.

In the second step, for each type of household in our data, we estimate how preferences for any given neighborhood would change if the level of rent and/or the economic and demographic composition of the neighborhood were to change. This step requires an instrumental-variables approach as the level of rent and the economic and demographic makeup of neighborhoods are likely jointly endogenously determined. We use two sets of instruments. Our first set of instruments are similar to those used in Bayer, Ferreira, and McMillan (2007) and Davis, Gregory, Hartley, and Tan (2017), neighborhood characteristics of nearby neighborhoods. Conceptually, these instruments help identify how changes in the level of rents affect preferences for any given neighborhood. For our second set of instruments, we exploit the U.S. Department of Housing and Urban Development's rules for designating Qualifying Census Tracts for Low Income Housing Tax Credits. As noted by Baum-Snow and Marion (2009), Diamond and McQuade (2017) and others, this rule creates a discontinuity based on tract-level poverty rates and median income. When we combine this discontinuity with the implications of our model, we are able to estimate each type's preferences over (a) the percentage of African-Americans ("black share") in the neighborhood and (b) the share of low-income households, defined as households earning in the bottom-tercile of income, in the neighborhood. We find that preferences for the demographic and economic composition of neighborhoods varies widely across our 315 types.

In the last section of the paper, we run counterfactual simulations of the model to understand the long-run impacts of (a) neighborhood composition and rent and (b) adult earnings of children to various housing policies, many of which have not been implemented. We compare the steady-state predicted allocations of people to neighborhoods and rents before the policy is implemented (and assuming people do not expect any changes) to the steady-state allocations and rents after the policy is implemented. This section highlights the importance of the structural approach to understand the impact of housing policies on long-run outcomes. Migration is costly and therefore households respond slowly to policy change. It may take many years to settle to a new steady state, and the immediate impact of the policy may look nothing like the new long-run steady state because of the process of slow migration.

In terms of understanding how low-income housing affects the socio-economic and demographic composition of neighborhoods, we find that the details of low-income housing policy matter quite a bit for the long-run outcomes. If policy makers introduce 100 new low-income housing units into only one tract in a metro area – roughly a 4 percent increase in the total number of housing units in that tract – the median impact is a 7.7 decline reduction in rent in that tract and a 0.5 and 1.7 percenage point increase in the black share and low-income share of residents living in the tract's existing housing stock. Similar to Diamond and McQuade (2017) we find larger rent declines when developments are placed in affluent neighborhoods, but there is substantial variation in the predicted impact of developments even after conditioning on pre-development neighborhood poverty rates and demographics. Ultimately, the outcome depends on the distribution of preferences over neighbors' race and income for the types of people likely to live in that tract; whether that tract provides a high level of intrinsic utility for many types of people; and, if there are close substitutes to that tract in the same metro area. A very different story emerges if policy-makers introduce 10 new low-income housing units to a given tract and to all the geographically proximate tracts until 10 percent of the tracts of the metro area have additional low-income housing. In this scenario, we find relatively little resorting by incumbent households in response to the policy and rent reductions are modest. This result appears roughly constant across tracts. The bottom line is that the introduction of a relatively large number of low-income housing units to a single tract has a high variance of possible outcomes; and the introduction of a relatively small number of low-income housing units to a large set of geographically proximate tracts induces a small change with very low variance.

In the last part of the paper, we use data from the recently released Opportunity Atlas of Chetty, Friedman, Hendren, Jones, and Porter (2018) on how Census tracts affect the later earnings of children (all else equal) to simulate the impact of a widescale expansion of Low Income Housing Tax Credits on the adult earnings of children moving into and out of tracts each receiving 100 new low-income housing units. We consider two cases, one in which the Opportunity-Atlast estimates of neighborhoods on adult earnings is fixed and another in which the equilibrium change in neighborhood composition can change the Opportunity-Atlas estimatess. Simulations show that if tracts receiving low-income units are placed randomly throughout a metro area, then the average impact of the program on earnings of children moving into and out of the tracts receiving the new units will likely be modest and negative, but with a large range of possible results. If policy-makers limit placement of new low-income housing units to the top third or so of potential locations, the improvement to total annual adult earnings of children as a result of the additional units is nearly $200,000 in the medium-sized MSAs that we study. We interpret these results as suggestive that a large-scale expansion of low-income housing tax credit policies to high-Opportunity-Atlas neighborhoods can positively impact the aggregate adult earnings of children, even after accounting for the possibility that the equilibrium re-sorting of the population affects the Opportunity Atlas estimates.

2 The Qualifying Census Tract Designation

A key feature of our paper is that we estimate household preferences over the socio-economic and demographic composition of their neighbors, enabling us to predict the frequency with which a given household will move to a different neighborhood if low-income housing is added to the neighborhood and the racial or economic composition changes as a result. Of course, households sort endogenously into neighborhoods based on observed and unobserved factors. Neighborhood demographic composition will therefore be correlated with unobservable factors, making estimation of preferences for demographic composition challenging. Our strategy to overcome this endogeneity problem is based on a Regression Discontinuity (RD) approach that exploits the discontinuous rule used by HUD to determine Qualifying Census Tracts (QCT) under the Low Income Housing Tax Credit (LIHTC) program. This discontinuous assignment rule generates variation in tract QCT status that is plausibly orthogonal to unobserved tract attributes. We will show that this exogenous variation in tract QCT status will provide the exogenous variation in neighborhood demographics needed for estimation of the structural model if 1) QCT affects the supply of low income housing, 2) QCT status affects the demographic composition of households moving into the neighborhood due to heterogeneity in preferences for nearby LIHTC developments, and 3) the nature of this demographic response varies across tracts according to initial demographic mix. As a starting point, this section provides RD estimates documenting each of these patterns in the data.

Each decade, the department of Housing and Urban Development (HUD) classifies some Census tracts as QCT for Low Income Housing Tax Credits (LIHTC) based on whether one of two conditions is satisfied according to data from the most recent Decennial Census: Tract median income is below 60% of the area median income, or, tract poverty rate is above 25%.[1] We study the impact of HUD's 2004 QCT designations, which were based on poverty

[1]LIHTC provides tax credits of up to 30% of the a development's property value. To receive a LIHTC credit, a developer must agree to set aside at least 20% of units in the development for individuals whose income is less than 50% of the area median gross income or set aside at least 40% of units in the development for individuals whose income is less than 60% of the area median gross income. Developers applying to the program submit proposals known Qualified Action Plans (QAP). These QAPs are scored by the local State

rates and median income from the 2000 Decennial Census.[2] Note that the QCT designation is one of two ways a neighborhood can be eligible for LIHTC credits.[3]

We verify that QCT status impacts the amount of low-income housing development. To cleanly show this point, we collapse tract poverty and the tract median income index to a one-dimensional "running variable,"

$$\mathcal{X}_j = max(Poverty_j - 0.25, 0.6 - MedIncIndex_j) \tag{1}$$

The QCT eligibility cutoff falls at $\mathcal{X}_j = 0$. Figure 1 shows the relationship of the probability a tract is designated as QCT to the value of the running variable. The figure shows that the probability a tract is designated as QCT jumps when the running variable hits 0, from a value of about 0.3 to a value of about 0.7.[4]

We now show how a set of tract outcomes Y_j vary with respect to the value of the running variable with regressions of the form,

$$Y_j \quad = \quad \beta_0 + \beta_1 1_{\mathcal{X}_j \geq 0} + g\left(\mathcal{X}_j\right) + \epsilon_j \tag{2}$$

$1_{\mathcal{X}_j \geq 0}$ is a dummy variable that is equal to 0 when $\mathcal{X}_j < 0$ and is equal to 1 when $\mathcal{X}_j \geq 0$ and $g\left(\mathcal{X}_j\right)$ is a 2nd-order polynomial in $\mathcal{X}$ that is allowed to have different coefficients when $\mathcal{X}_j$ is above and below 0. The parameter of interest is β_1, which measures the jump in the conditional expectation of the outcome variable when the running variable is at least 0, i.e. when the threshold for QCT status is achieved.

Figure 2 shows the expected value of Y_j as a function of $\mathcal{X}_j$, where Y_j refers to the building of any new low-income units (left), at least 30 new low-income units (center) and at least 100 new low-income units (right). Our sample includes all tracts in a metro area in the United States in the year 2000. Our tract-level data are for the cumulative number of new

Housing Finance Agency on an annual basis, and awards are made to the highest scoring applicants until funds are exhausted.

[2] Baum-Snow and Marion (2009) exploit the median-income cutoff for eligibility for LIHTC in the 1990s (based on the 1990 Census) to estimate the program's impact on a host of neighborhood-level outcomes. For our purposes, this does not yield sufficient statistical power to evaluate the impact of LIHTC in the 2000s. (Note that in 2000 and earlier, QCT status was recalculated following each decennial Census, and is now regularly recalculated based on measures from the American Community Survey). The likely explanation is that, as shown in Figure 3 relatively few neighborhoods fall close to the median income threshold for QCT designation. Many more tracts fall close to the poverty rate threshold, and as we show below, exploiting the full two-dimensional threshold in the RDD yields sufficient power to detect program impacts.

[3] A tract is also eligible for LIHTC credits if HUD designates it as a Difficult Development Area (DDA), defined as having a ratio of construction costs to area income above a particular threshold.

[4] In this section, we only include tracts where $-0.2 \leq \mathcal{X}_j \leq 0.2$. Later on in the paper, when we explicitly use QCT status to identify preference parameters, we include all tracts in the analysis and specify a more flexible functional form for the control function of poverty rates and income.

low-income units from 2004 until 2013 and are computed from the HUD LIHTC database. In each panel there is a clearly visible jump that occurs when $\mathcal{X}_j$ is zero.

Our RD strategy for estimating the effects of QCT status relies on the assumption that, while unobserved confounding factors will differ between QCT and non-QCT on average, QCT status is as good as randomly assigned for tracts with income/poverty pairs close to the QCT cutoff (Baum-Snow and Marion (2009)) if the distribution of unobserved amenities changes smoothly as a function of median income and poverty. This assumption of "unconfoundedness" across the cutoff is not directly testable, but we perform falsification exercises that are standard in the RD literature to check for observable patterns that question this assumption. Figure 3 plots the QCT cutoff line against income and poverty rates; the figures shows no evidence of bunching at the eligibility boundary which, if present, is commonly interpreted as evidence of non-random manipulation of the running variable(s). Additionaly, Table 1 presents balance tests for tract variables from the 2000 Census, which were pre-determined in 2004 when QCT status was designated. With the exception of family income, we find no statistically significant differences in the values of these observable tract characteristics above versus below the value $\mathcal{X}_j = 0$.

Finally, we show how the probability that black, hispanic, and low-income households move to a tract varies with that tract's value of $\mathcal{X}_j$. In this analysis, we use information on location choices from a large, household-level annual panel data set on location decisions. We discuss this data in great detail in section 4, but for now we note that the data are from the NYFRB/Equifax Consumer Credit Panel. For the analysis in this section, we only include households that move to a different Census tract after 2004 and we only include data in the years of a move.

Table 2 shows the change in the *rescaled* probability that a household moves to a tract at $\mathcal{X}_j = 0$ by various demographic and economic characteristics of the household: race (black, hispanic, white and "other"), income ("low income" and "non-low income") and age of the household when they first appear in the NYFRB/Equifax data set (Less than 35, 35-44, 45-54, 55-64, and 65 and over).[5] The column marked "All Neighborhoods" shows the impact on the rescaled probability for all the neighborhoods in the sample; the remaining three columns show the impact for, respectively, majority black, hispanic, and white or other neighborhoods as measured from tract-level data in the 2000 Census.[6] We set rescaled probabilities equal to

[5] The NYFRB/Equifax data do not include information on income or race. As we discuss later, we sort households into types. We identify the average income for each type via a regression of tract-level income on shares of types in each tract and we identify race using information on the Census block in which the household is first observed – see footnotes 12 and 19 for more details.

[6] There are three groups of households in the 2000 Census: Black, Hispanic and White/Other. We characterize a tract based on the largest racial/ethnic group in that tract.

empirical probabilities multiplied by the number of tracts in the metro area of the household of residence. One convenient way to interpret the values in table 2 is they are equal to 100 times the change in probability that a person moves to a tract at $\mathcal{X}_j = 0$ as compared a tract where $\mathcal{X}_j$ is slightly less than zero in a metro area with 100 tracts. Figure 4 shows results graphically; the y-axis shows the rescaled probability a given tract is chosen and the x-axis is $\mathcal{X}_j$. In table 2 and figure 4, we rescale the raw probabilities of the households in our sample because these households live in metro areas of vastly different sizes, and pooling raw probabilities across these households will, roughly speaking, relatively overweight the experiences of people in small metro areas and underweight the experiences of people in large metro areas.

Table 2 shows all results and Figure 4 visualizes results we discuss here. First, when considering all tracts, all black households and low-income hispanic households are more likely to move into a tract with $\mathcal{X}_j$ just above 0 as compared to a tract with $\mathcal{X}_j$ just below 0. The impact on choice probabilities of $\mathcal{X}_j$ crossing the 0 threshold is 0.151 percentage points for black-households and 0.105 percentage points for hispanic low-income households. This is in contrast to white and "other" households, who are not more likely to move when $\mathcal{X}_j$ becomes 0. In sum, the data show that the when $\mathcal{X}_j$ crosses 0, and the probability that a tract is designated as QCT jumps, this systematically and differentially affects the neighborhood choice probabilities of some black, hispanic, and low-income households.

3 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 should live. As in Kennan and Walker (2011) and Bayer, McMillan, Murphy, and Timmins (2015), we model location choices in a dynamic discrete choice setting. For purposes of exposition, we write down the model describing the optimal decision problem of a single family which enables us to keep notation relatively clean. For now, we consider a model of within-MSA location choices, and estimate separate models for each MSA.[7] When we estimate the parameters of this model, we will allow for the existence of many different "types" of people in the data. Each type of person will face the same decision problem, but the vector of parameters that determines payoffs and choice probabilities will be allowed to vary across types of people.

The decision problem of the household is very similar to the one described within a partial equilibrium framework in Davis, Gregory, Hartley, and Tan (2017). The family can choose

[7] A straightforward extension would nest these models inside a model of MSA choice.

to live in one of J locations. Denote j as the family's current location. We write the value to the family of moving to location ℓ given a current location of j and current value of a shock ϵ_ℓ (to be explained later) as

$$V(\ell \mid j, \epsilon_\ell) \;=\; u(\ell \mid j, \epsilon_\ell) + \beta EV(\ell)$$

In the above equation $EV(\ell)$ is the expected future value of having chosen to live in ℓ today and β is the factor by which future utility is discounted. We assume the household problem does not change over time, explaining the lack of time subscripts.

u is the flow utility the agent receives today from choosing to live in ℓ given a current location of j and a value for ϵ_ℓ. We assume u is the simple function

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

δ_ℓ is the flow utility the household receives this period from living in neighborhood ℓ, inclusive of tastes for rents, neighborhood demographics, and any amenities or natural advantages the neighborhood provides; κ are the fixed costs (utility and financial) a household 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 ϵ_ℓ is a random shock that is known at the time of the location choice. ϵ_ℓ is assumed to be iid across locations, time and people. The parameters δ_ℓ and κ may vary across households, but for any given household these parameters are assumed fixed over time. ϵ_ℓ induces otherwise identical households living at the same location to optimally choose different future locations.

Denote ϵ_1 as the shock associated with location 1, ϵ_2 as the shock with location 2, and so on. In each period after the vector of ϵ are revealed (one for each location), households choose the location that yields the maximal value

$$V(j \mid \epsilon_1, \epsilon_2, \ldots, \epsilon_J) \;=\; \max_{\ell \in 1,\ldots,J} V(\ell \mid j, \epsilon_\ell) \tag{3}$$

$EV(j)$ is the expected value of (3), where the expectation is taken with respect to the vector of ϵ.

While this model looks simplistic, it is the workhorse model used to study location choice. Differences in models reflect specific areas of study and availability of data. For example, in their study of migration across states, Kennan and Walker (2011) replace δ with wages after adjusting for cost of living.[8] Bishop and Murphy (2011) and Bayer, McMillan, Murphy,

[8]In our model and that of Kennan and Walker (2011), the only choice households make is where to live each period. To be clear, there are many differences between the two models in the state space, expected

and Timmins (2015) specify δ as a linear function of spatially-varying amenities with the aim of recovering individuals' willingness to pay for those amenities. We allow the δ's to vary flexibly across neighborhoods, with the aim of realistically forecasting the substitution patterns that are likely to occur in response to government policies that change the relative prices of neighborhoods.

When the ϵ are assumed to be drawn i.i.d. from the Type 1 Extreme Value Distribution, the expected value function $EV(j)$ has the functional form

$$EV(j) \quad = \quad \log\left\{\sum_{\ell=1}^{J} \exp \widetilde{V}(\ell \mid j)\right\} + \zeta \tag{4}$$

where ζ is equal to Euler's constant and

$$\widetilde{V}(\ell \mid j) \quad = \quad \delta_\ell \; - \; \kappa \cdot 1_{\ell \neq j} \; + \; \beta EV(\ell) \tag{5}$$

That is, the tilde symbol signifies that the shock ϵ_ℓ 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 $\delta_1, \ldots, \delta_J$, κ and β and of observed choice probabilities. In other words, a likelihood over choice probabilities observed in data can be generated without solving for value functions.[9]

4 Data and Likelihood

Like Davis, Gregory, Hartley, and Tan (2017), 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 5% sample.[10] For years 1999 to 2014, 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 race, education,

utility associated with each location, and how costs vary with specific moves.

[9]See Davis, Gregory, Hartley, and Tan (2017) for more details.

[10]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. Note that a SSN is also required to receive a housing voucher.

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. There are no other sample restrictions.[11] 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 145,421,128.

We stratify households into types using an 8-step stratification procedure. Note that when we assign households to types we use no information on location. We begin with the full sample, and subdivide the sample into smaller "cells" based on (in this order): The racial plurality, as measured by the 2000 Census, of the 2000 Census block of residence (4 bins),[12] 5 age categories (cutoffs at 30, 45, 55, and 65),[13] number of adults age 18 and older in the household (1, 2, 3, 4+), and then the presence of an auto loan, credit card, student loan and consumer finance loan. We do not subdivide cells in cases where doing so would result in at least one new smaller cell with fewer than 250,000 observations. In a final step applied to all bins, we split each bin into three equally-populated types based on within-bin credit-score terciles. After all the dust settles, this procedure yields more than 315 types of households.

The number of Census tracts varies by MSA, and as mentioned we estimate preferences by types separately for each MSA.[14] Allowing a separate value of δ for each tract and for each type would require estimating more parameters than is feasible given the size of our data. Therefore, for parsimony, and to exploit the fact that geographically nearby tracts likely provide similar utility, for each type we specify that the utility of location j, δ_j, is a

[11] Davis, Gregory, Hartley, and Tan (2017) restrict the sample to renters, but we include all households – renters and owners – in our sample.

[12] We assign race based on the racial plurality of all persons in the Census block, owners and renters. We expect that the geography of the Census block is small enough that the racial plurality of renters will be identical to that of the entire block. We classify individuals based on the racial plurality of the block where they are first observed, which in most cases is 1999.

[13] Whenever we refer to a household "age" in the FRBNY Consumer Credit Panel / Equifax data, we are referring to the age of the person in the household in the initial random sample. We are not using the ages of any other people in the household.

[14] In the case of Los Angeles, Davis, Gregory, Hartley, and Tan (2017) consider preferences over 1,748 Census tracts.

function of latitude (lat_j) and longitude (lon_j) of that location according to the formula

$$\delta_j = \sum_{k=1}^{K} a_k B_k \left(lat_j, lon_j \right)$$

The B_k are parameter-less basis functions. For each type and for each MSA, we use $K = 100$ basis functions. Inclusive of the moving cost parameter, we estimate $100+1 = 101$ parameters per type. With more than 300 types, we estimate more than 30,000 parameters.

To define the log likelihood that we maximize we need to introduce some more notation. Let i denote a given household, t a given year in the sample, j_{it} as person $i's$ starting location in year t and ℓ_{it} as person $i's$ observed choice of location in year t. Denote τ as type and the vector of parameters to be estimated for each type as θ_τ. The log likelihood of the sample is

$$\sum_{\tau} \sum_{i \in \tau} \sum_{t} \quad p \left(\ell_{it} \mid j_{it}; \theta_\tau \right) \tag{6}$$

$p(.)$ is the model predicted log-probability of choosing ℓ_{it} given j_{it}. For each τ we use the quasi-Newton BFGS procedure to find the vector θ_τ that maximizes the sample log likelihood.

The likelihood considers all within MSA moves. The likelihood excludes any moves to or from an MSA.

5 Preferences for Neighborhood Composition and Rent

5.1 Specification of Utility

We specify that the utility that type τ receives of living in neighborhood j, $\delta_{j\tau}$, is a function of the log of rent that is paid, the share of black households that live in the neighborhood, the share of low-income households that live in the neighborhood and amenities in neighborhood j that are unobservable to us.

$$\delta_{j\tau} = \bar{\delta}_{j\tau} + \alpha_{R\tau} R_j + \alpha_{B\tau} B_j + \alpha_{L\tau} L_j + \alpha_{A\tau} A_j \tag{7}$$

In equation (7), log rent paid in neighborhood j is R_j, the share of black households in neighborhood j is B_j, the share of low-income households in neighborhood j is L_j and unobservable amenities in neighborhood j are A_j. $\alpha_{R\tau}$, $\alpha_{B\tau}$, $\alpha_{L\tau}$ and $\alpha_{A\tau}$ reflect type τ preferences for rent, black share, low-income share and amenities, respectively. $\bar{\delta}_{j\tau}$ is the normalized level of utility when all the other variables are equal to 0.

Denote our maximum likeilhood estimate of $\delta_{j\tau}$ from the previous section as $\widetilde{\delta}_{j\tau}$. We do not regress $\widetilde{\delta}_{j\tau}$ on rent paid, black share and low-income share as this would yield biased estimates of the coefficients: Amenities are unobserved and we expect amenities to be correlated with all these variables.[15] Therefore, we use an instrumental variables approach to estimate the type-specific coefficients $\alpha_{R\tau}$, $\alpha_{B\tau}$ and $\alpha_{L\tau}$. To be valid, these instruments must be correlated with the endogenous variables R_j, B_j, and L_j (relevant) and uncorrelated with unobserved neighborhood amenities A_j (exogenous).

5.2 The Bayer et. al. Instruments

One set of instrumental variables we use is a vector of housing stock characteristics for homes located between 5 and 20 miles from the neighborhood; these variables are labeled as $\mathcal{H}_j$. This instrument is in the spirit of Bayer, Ferreira, and McMillan (2007) and Davis, Gregory, Hartley, and Tan (2017). These instruments are *relevant* (i.e. correlated with R_j) because characteristics of potential substitutes for a neighborhood should affect its equilibrium rental price. These instruments are *exogenous* (i.e. uncorrelated with A_j) under the assumption that characteristics of places sufficiently far from j have no direct effect on j's amenity valuation. We also include characteristics of the neighborhood's own housing stock and housing stock characteristics for homes located between 0 and 5 miles from tract j, labeled as $\mathcal{O}_j$, as controls in regressions that follow. Our tract-level data on $\mathcal{H}_j$ and $\mathcal{O}_j$ are from the U.S. Census Bureau from the year 2000.

5.3 Low Income Housing Tax Credit Eligibility

We construct a second set of instruments for B_j and L_j based on an exogenous source of variation in tract eligibility for Low Income Housing Tax Credits (LIHTC), combined with tract-specific predictions from our model about the likely impact of eligibility on each tract's demographic mix. The logic behind the IV strategy resembles the "shift-share" IV approaches of Card (2001) to study the labor market impacts of immigration and Boustan (2010) to study white flight in response to black migration to northern U.S. cities, but our approach uses the model to combine information on types' responses to exogenous changes in one particular neighborhood amenity with information on the type mix of households who are marginal in their location choices.

For the instruments to be a valid, they must be correlated with the demographic measures B_j and L_j and uncorrelated with unobserved amenities A_j. The instruments also must vary

[15] Rent will obviously be correlated with unobserved amenities. Additionally, as long as different types of households have different values for $\alpha_{A\tau}$ i.e. different preferences for amenities, type shares by neighborhood will be correlated with unobserved amenities. Since since types vary by race and income, the black share and low-income share of each neighborhood will also be correlated with unobserved amenities.

independently despite being constructed from one source of variation in tract eligibility for the policy. As discussed in section 2, we use the plausibly exogenous variation in QCT status to generate variables that are correlated with tracts' black and low-income shares (B_j and L_j) but are uncorrelated with other amenities (A_j). Figure 4 and table 2 show that black and low-income households are more likely to move to tracts likely to be designated as QCT, that is with running variable $\mathcal{X}_j$ greater than or equal to zero as compared to less than zero. This suggests some types' indirect utility is influenced by a tract's QCT status, a nessary condition for our procedure to create instruments.

Note that we cannot take the "direct" approach of regressing black-share and low-income share on QCT status, generating predicted values, and then regressing our maximum likelihood estimate of $\widetilde{\delta}_{j\tau}$ on these predicted values. The reason is that the predicted values of black-share and low-income share would be co-linear, low when $Q_j = 0$ and high when $Q_j = 1$. Instead, we use a four-step procedure where we use the predictions of our decision model to provide independent variation of black share as compared to the low-income share.

In the first step, we regress by 2SLS of our maximum-likelihood estimate of type-specific neighborhood preferences on log-rent, QCT status, other observables and a spline of poverty rates and median income:[16]

$$\widetilde{\delta}_{j\tau} \quad = \quad d_{0\tau} + d_{1\tau} R_j + d_{2\tau} Q_j + d_{3\tau} \mathcal{O}_j + g\left(pov_j, inc_j; d_{4,\tau}\right) + v_{j\tau} \tag{8}$$

We instrument for QCT status Q_j and log-rent R_j using the following specification for each, and allowing estimated coefficients to vary:

$$E_j \quad = \quad f\left(pov_j, inc_j\right) I_j + b_1 \mathcal{H}_j + b_2 \mathcal{O}_j + g\left(pov_j, inc_j; b_4\right) + e_j \tag{9}$$

where E_j is the endogenous variable, either Q_j or R_j, and e_j is the error which, of course, varies with the endogenous variable. The regressors include the dummy variable I_j that is equal to 1 if tract poverty is above 25% or median income is below 60% of area median income and 0. The function f allows the probability of the QCT designation to jump by different amounts at different parts of the income and poverty border for QCT eligibility. $\mathcal{H}_j$ includes characteristics of the housing stock 5-20 miles away (an instrument). Since the cubic spline controls for smooth changes in the expectation of the dependent variable as a function of tract income and poverty rates, identification of the impact of QCT_j on Y_j, β_2,

[16]To reduce complications arising from sampling variability in our estimates of type-specific indirect utilities (which are estimates from a non-linear model), we restrict the estimation sample to include only the type-specific indirect utility estimates coming from MSAs where the type in question is observed at least 3000 times. This reduces the number of micro-level observations underlying the procedure from 145,421,128 to 105,048,992.

relies on the discontinuous jump in QCT status at the border, and not comparisons of tracts with income and poverty rates far from the QCT eligibility cutoff.

Denote the predicted values arising from the estimated coefficients in (8) as $\widehat{\delta}_{j\tau}$. These predicted values will *jump* at the QCT boundary holding all other observables constant.[17] Since the change in QCT status is orthogonal to amenities, which are assumed to be smooth through the QCT border, variation in $\widehat{\delta}_{j\tau}$ induced by changes to QCT status will be unrelated to changes in unobserved amenities.

Figure 5 plots the distribution of the t-statistic across types associated with the null that $d_{2\tau} = 0$. This figure demonstrates that types differentially care about QCT status. It is clear that a large majority of types have preferences for QCT eligibility that are significantly different from zero, but the distribution is bimodal. A majority of types place a negative valuation of QCT-eligibility, but substantial minority of types place a positive valuation on QCT-eligibility.

In the second step, we solve and simulate the model for every type using $\widehat{\delta}_{j\tau}$ as the flow payoff for location j for type τ before factoring in any moving costs.[18] Since we know the initial distribution of types across locations, and we know the model-implied probability each type moves to location j from any starting location, we simulate the steady-state distribution of types in all locations. Denote $\widehat{s}_{j\tau}$ as the simulated steady-state percentage of neighborhood j accounted for by type τ arising from this step.

In the third step, we map our distribution of simulated steady-state types in each location, $\widehat{s}_{j\tau}$ into a simulated black share $\widehat{B}_j$ and simulated low-income share $\widehat{L}_j$. We can make this mapping because associated with each type τ is a race (Black, Hispanic, etc.) and an income level. Once we know the distribution of types, we can figure out the share of the neighborhood that is black and the share with income in the first tercile.[19]

This procedure creates independent variation in the instruments $\hat{B}_j$ and $\hat{L}_j$ for two reasons. First, the $d_{2\tau}$ term varies across types. In other words, types vary in how much they care about QCT, consistent with the reduced-form results of table 2 and figure 4. Second, there is variation across neighborhoods in what mix of types are marginal with respect to moving in or out, which is largely determined by (a) type-specific differences in the other coefficients in equation (8) and (b) variation across MSAs in proportion of the total population accounted for by the different types.

[17]The size of the jump will depend on the type-specific value of $d_{2\tau}$.

[18]In other words, we replace our maximum likelihood estimates of $\widetilde{\delta}$ with these $\widehat{\delta}$, but keep all other estimated model parameters identical when simulating the model.

[19]We determine the types in the first income tercile by regressing log income, measured at the tract level, against tract-shares by type. The estimated coefficient on the type share variable is an average-income index. Given this index, we can pick out the types constituting the bottom-third of the income quintile.

In the fourth and final step we estimate the parameters of interest in the utility function, $\alpha_{R\tau}$, $\alpha_{B\tau}$ and $\alpha_{L\tau}$, that we need to implement counterfactual simulations. To do this, we estiamte a 2SLS regression of $\widetilde{\delta}_{j\tau}$ on rent, black share, low-income share, QCT status, other tract-level observables, and the spline in poverty and income.

$$\begin{aligned}\widetilde{\delta}_{j\tau} &= \\ &\bar{\delta}_{j\tau} + \alpha_{R\tau} R_j + \alpha_{B\tau} B_j + \alpha_{L\tau} L_j + +a_{1\tau} Q_j + a_{2\tau} \mathcal{O}_j + g\left(pov_j, inc_j; a_{3\tau}\right) + \nu_{j\tau}\end{aligned} \tag{10}$$

where we instrument for the endogenous variables $E_j = \{R_j, Q_j, B_j, L_j\}$ using the specification in equation (9) but adding the generated variables $\widehat{B}_j$ and $\widehat{L}_j$ as additional instruments.

Figure 6 plots the distribution across types of the t-statistic for $\alpha_{R\tau}$, $\alpha_{B\tau}$ and $\alpha_{L\tau}$. The vast majority of types' (92%) have a negative preference for rent. For a majority of types (79%), a tract's black share negatively affects indirect utility. The impact of the first-income tercile share on indirect utility also appears to be bimodal.

6 Long-Run Impact of Low-Income Housing Units

We are interested in the long-run impact of additional development of low-income housing units on the rent, black-share and low-income share of the Census tract in which the development occurs. To do this, we compare simulated steady-states of the model for a baseline case, where households assume no changes to policy, and a "counterfactual" case, the steady state of the model after the low-income housing units are built.

In the first set of simulations that we call the "first policy," we assume either 10, 50, 100 or 250 low-income units are built in a targeted tract, but no other low-income units are built. A Census tract contains approximately 2,500 units, so development of an additional 100 units increases the total stock of units in the tract by about 4%. We repeat this in a seperate experiment for every tract in all MSAs with at least 100 and no more than 250 tracts. The population of the 46 MSAs in this sample ranges from 385 thousand (Beaumont-Port Arthur, TX) to 1.23 million (Nashville, TN). We assume the newly built low-income units are populated with new residents to the MSA, implying no existing residents need to move. The households living in the newly built low-income housing units are assumed to earn income in the bottom income tercile; additionally, we assume the black share of these households is equal to the black share of households in the bottom income tercile in that MSA. In each simulation, we keep track of the changes between the counterfactual and baseline in rent, black-share and low-income share in the tract receiving the units.

In the second set of simulations, our "second policy," we assume 10 low-income units are built in one tract, call it the "target" tract, *and* 10 low-income units are built in each of the nearest 10% of all tracts of the metro area.[20] As before, we assume the new units are populated with low-income residents that are new to the MSA. This experiment removes incentives for some households to relocate for two reasons: (1) the number of additional low-income units in any given tract is relatively small and (2) proximate tracts also have additional low-income housing units, removing much of the ability for households to move to similar tracts without the additional low-income units. We repeat these simulations for the same tracts as in the first policy simulations, and in each simulation we measure the changes in rent, black-share and low-income share in the target tract (but not the surrounding tracts) relative to baseline.

Before moving on to results, we need to discuss how we compute steady states, and to do that we must first discuss utility. We assume the utility type τ receives from living in tract j in the baseline is $\widetilde{\delta}_{j\tau}$. Denote the log rent, black share, and low-income share in tract j in the baseline steady-state as R_j^b, B_j^b and L_j^b and in the counterfactual steady-state as R_j^c, B_j^c and L_j^c. We compute steady-state utility for type τ from living tract j in the counterfactual as

$$\widetilde{\delta}_{j\tau} \ + \ \alpha_{R\tau}\left(R_j^c - R_j^b\right) \ + \ \alpha_{B\tau}\left(B_j^c - B_j^b\right) \ + \ \alpha_{L\tau}\left(L_j^c - L_j^b\right) \tag{11}$$

Next, we define a steady-state equilibrium in our model. In a steady-state equilibrium, in every neighborhood j in a given metropolitan area (1) the expected share of black residents in the tract is equal to the actual share, (2) the expected share of low-income residents is equal to the actual share, (3) housing demand is equal to the number of available units, (4) rents are stable and (5) the population and the shares of black residents and low-income residents are fixed. We compute these shares and the population as follows. Define the measure of type-τ households living in tract ℓ as $\eta_\tau(\ell)$ and let $\mathcal{I}_{\tau B}$ and $\mathcal{I}_{\tau L}$ denote indicator functions for whether the type-τ household is black or low-income.[21] Denote the probability that type τ moves to tract j while living in tract ℓ in the baseline simulations as $\rho_{j\tau}^b(\ell)$.[22] We can compute the total number of housing units demanded (H_j^b), the black share and the

[20]Distance from the target tract is measured from tract centroid to tract centroid. Conceptually the experiment adds 10 low-income units for all tracts in a circle around the target tract, assuming the target tract is not at or near the metro boundary.

[21]We assume low-income is a permanent attribute of the household.

[22]Of course ℓ can equal j, in which case $\rho_{j\tau}^b(\ell)$ is the probability the household does not move.

low-income share in tract j in the baseline as:

$$H_j^b = \sum_{\tau}\sum_{\ell}\eta_\tau(\ell)\,\rho_{j\tau}^b(\ell) \tag{12}$$

$$B_j^b = \left(H_j^b\right)^{-1}\sum_{\tau}\mathcal{I}_{\tau B}\left[\sum_{\ell}\eta_\tau(\ell)\,\rho_{j\tau}^b(\ell)\right] \tag{13}$$

$$L_j^b = \left(H_j^b\right)^{-1}\sum_{\tau}\mathcal{I}_{\tau L}\left[\sum_{\ell}\eta_\tau(\ell)\,\rho_{j\tau}^b(\ell)\right] \tag{14}$$

If we define $\rho_{j\tau}^c(\ell)$ as the probability type τ moves to j from tract ℓ in the counterfactual, we compute H_j^c, B_j^c and L_j^c analogously.

To compute housing supply in each tract, in the baseline we set rents equal to those in the data and then find the housing supply such that steady-state housing demand, as defined in equation (12), is equal to housing supply. To find housing supply in the counterfactual steady-state, we need to take a stand on how housing supply varies with prices. Keeping notation consistent with earlier sections, denote H_j^b and H_j^c as the log of the steady-state number of housing units of tract j in the baseline and counterfactual simulations; and denote the market-clearing log rent in tract j in the baseline and in the counterfactual as R_j^b and R_j^c. We specify

$$\ln H_j^c - \ln H_j^b = \psi\left(R_j^c - R_j^b\right). \tag{15}$$

We set $\psi = 0.25$ in all simulations.

Finally, we need to discuss the possibility the model admits multiple steady-state equilibria and how we react to that possibility. Models where people have preferences over attributes of their neighbors sometimes allow for multiple steady states. To understand why, consider a framework where (i) there are two types of people, black and white, and two tracts A and B; (ii) both white and black types prefer to live in perfectly segregated tracts; and (iii) neither white nor black people have any intrinsic attachment to either tracts A or B. Two equilibria likely exist in this environment, a first where white types live in A and black types live in B, and a second where black types live in A and white types live in B.

Using similar reasoning, we believe our model may admit multiple equilibria. We take steps to keep the steady-state equilibrium we compute in the counterfactual as "close" as possible to the baseline equilibrium and to avoid reporting large changes simply due to the possibility of multiplicity. Using the example outlined in the previous paragraph, if whites occupy tract A and blacks occupy tract B in the baseline, we attempt to avoid computing equilibria in the counterfactual where whites occupy B and blacks occupy A and nothing else has changed. To compute the counterfactual steady state, along the computational transition path from the baseline to the new steady state we assume agents have backward-looking expectations about the values of R_j, B_j and L_j in each tract. Every change in those variables is a shock to households in the model: after each period along the transition path we assume (only for the purposes of computing the new steady state) that households assume those variables remain fixed forever. This anchors the new steady-state to the old steady-state, at least for the first few periods along the computational transition path. The hope is that by forcing households to look backward, we eliminate any large jumps in the black- or low-income share along the computational transition path simply due to a change in the nature of expectations. Note that once the new steady state is computed, it is fully rational for households to have constant expectations for R_j, B_j and L_j.

We also Winsorize the top and bottom 10% of values of $\alpha_{R\tau}$, $\alpha_{B\tau}$ and $\alpha_{L\tau}$, i.e. since there are 315 types we replace the top and bottom 31 estimates of each parameter with the 32nd highest or lowest value. When we compute new steady states without Winsorizing, or with Winsorizing at a 5% level, we occasionally compute what appear to us to be implausable jumps in counterfactual steady-states. In these cases, the newly built low-income units lead to a new steady state in which there is a large increase in the tract's low-income share and in the level of rent. These changes are sustained in equilibrium because a large increase in the tract's new population is from low-income types with a strong preference for low-income neighbors. We try to rule out this kind of change in steady states because it seems similar to the change in steady-state equilibria in the example we describe earlier where whites that used to live in A now live in B and blacks that used to live in B now live in A.

6.1 Housing Units Built in Only One Tract

Figure 7 shows the impact on log rent (top panel), black share (middle panel) and low-income share (bottom panel) when 10, 50, 100 and 250 units are added to a single tract. We repeat this policy experiment for each of the 839 tracts in our sample, one tract at time; each blue dot represents the experience of the tract in which the low-income units are added. The solid blue line traces through the outcomes of the median tract, the identity of which can change.[23] The top panel shows that, when considered at the median tract, rents fall relative to the baseline as low income units are added. But, as the distribution of blue dots also shows, the range of outcomes is large. For example, when 50 units are added, rents may not decline at all or they may decline by nearly 15 percent.

[23]We exclude the top and bottom 5% of dots from the graph so the range of the y-axis is smaller, making it easier to see any patterns that emerge. The graphed median is inclusive of all results.

The middle panel shows how the black-share of each tract changes as low-income housing units are added to the tract. The y-axis reports changes to the black share and, in the panel below, the low-income share of *only* the existing housing units in the tract. If, for example, the low-income share of existing housing units does not change, then we can infer the tract has become more economically integrated, as the new low-income housing units are populated entirely by low-income residents. The middle panel shows that for most tracts, the percentage of black residents living in the existing housing stock does not change. The range of possible outcomes is quite small, increasing at most by 3 percentage points.

The bottom panel shows how the low-income share of each tract changes as low-income housing units are added to the tract. The panel shows that at the median, the percentage of low-income residents in existing units as the number of low-income units increases, although there is wide dispersion around the median: For example, in some tracts an increase of 50 or more low-income housing units boosts the low-income share of existing units by 9 percentage points, approximately 225 units.

In summary, the three panels tell the following story: For the median tract, when new low-income housing units are introduced the percentage of black residents in existing units stays constant, the percentage of low-income households in those units increases, and rents fall. However, there is considerable variation around these results. In some tracts, for example, the share of low-income residents living in existing units might fall. The wide range of possible results show in figure 7 suggest to us that diff-in-diff estimates on the impact of low-income housing relying on identification from a small number of tracts might produce a wide range of estimates.

We sliced the data a number of ways to see if we could infer more intuition about what variables were associated with the change in rent, black-share and low-income share at the tract level. For this analysis, we fixed at 100 the number of new low-income units in a given treated tract. Figure 8 shows the results when we cut the data by tract-level poverty rate, left column, and black share, right column; both the tract-level poverty rate and tract-level black share on the x-axis are taken from the 2000 Census.[24] The figures in this graph reinforce the idea that even after controlling for tract poverty rate or black share, there is a healthy range of possible outcomes for the change in rent or low-income share, and a smaller range of outcomes for the change in black share. Interestingly, the column on the left shows that the policy has the biggest impact (when measured at the median tract) on rents, the black share and the low-income share for tracts with a poverty rate of around 20%. Still, even at

[24]We also investigated relationships for MSA-wide variables on black share, a racial segregation index and an income segregation index. The results are not materially different from what we present next so we omit them.

a 20% poverty rate, many outcomes are possible, as shown by the dispersion of blue dots.

6.2 Housing Units Built in a Cluster of Tracts

In the second set of simulations we assume 10 low-income units are built in a "target" tract and in each of the proximate tracts until 10% of the tracts in the MSA have additional low-income housing units built. We chose this experiment to ask what happens if housing policy reduces incentives for households to move to nearby, similar tracts that may be close substitutes to the target tract. Given the MSAs in our sample have between 100 and 250 tracts, in our experiment we add in total between 100 and 250 new low-income housing units. As before, we assume the new low-income units are populated by people moving to the MSA and not existing residents. In the graphs below, we only report changes to the target tract; we do not include changes that may have occurred to the surrounding tracts. As with the first set of simulations, when we report changes to black-share or low-income share, we only report the changes to the incumbent housing units. Thus, when we report a 0 or no change, it implies the entire low-income population of the target tract increased by the amount of newly built low-income housing units.

Using the same format as in Figure 8, Figure 9 shows the results for the second set of simulations as red dots. Each red dot in the figure represents the experience of the target tract and the solid red line traces through the experience of the median tract for the x-axis. The blue dots in these figures are the same as in Figure 8.

What jumps out from the red dots in these figures is that in the second set of policy simulations, the black share and the low-income share of the existing housing units essentially does not change and rents decline very modestly. The red dots showing individual tract experiences are clustered very tightly around the solid red line depicting the median, suggesting conclusions we draw from the median are likely applicable to almost every tract. The types of households living in existing housing units in the target tract do not move; and rents in the target tract fall very slightly to ensure these types of households do not want to move. If the goal of policy-makers is to promote more racial and socio-economic integration, the second set of policy simulations suggest that adding a small number of low-income units to a relatively large number of tracts around a target tract achieves that goal, as new low-income residents move in and the types of residents in existing units do not change.

Tables 3, 4 and 5 show the results of regressions on the change in rent, the black-share and the low-income share induced by the additional low-income units.[25] Columns 1-4 show the results from the "policy 1" simulations where 100 units of low-income housing are added

[25]Note that the range of the black-share and poverty rate is $[0, 1]$ and not $[0, 100]$.

to one tract and columns 5-8 show the results from the "policy 2" simulations where 10 units are added to a target tract and each of 10% of the nearest tracts. Columns 1 and 5 show the results with only tract-level regressors; columns 2 and 6 show the results with only MSA-level regressors; columns 3 and 7 show the results with both tract- and MSA-level regressors, but without interactions; and columns 4 and 8 show the results with tract- and MSA-level regressors and interactions.

Before discussing results, we should define two MSA-level regressors, the Racial Dissimilarity Index and the Income Segregation Index. Denote b_j as the number of black households living in tract j; denote w_j as the number of white households living in tract j; and denote B and W as the black and white population of the MSA. The Racial Dissimilarity Index for an MSA is computed as

$$\frac{1}{2}\sum_j \mid \frac{b_j}{B} - \frac{w_j}{W} \mid$$

The index is equal to 0 when black and white households live in the same proportion relative to the MSA total in every tract. And the index is equal to 1 when tracts are perfectly segregated by race. The Income Segregation index is defined as

$$\frac{1}{J}\sum_j \frac{finc_j}{\overline{finc}} \times \ln\left(\frac{finc_j}{\overline{finc}}\right)$$

where $finc_j$ is the average family income of tract j in the 2000 decennial Census and $\overline{finc}$ is the MSAs average family income. The index is equal to zero if all tracts have the same average family income and increases as average family income across tracts diverges.

Focusing on results, looking across all the sets of regressions the MSA-level variables (columns 2 and 6) tend to better predict the change in tract-level outcomes in response to the new low-income housing units than the tract-level variables (columns 1 and 5). Columns 3 and 7 show regressions with with both tract-level and MSA-level variables. When low-income housing units are added, tracts located in MSAs that are more racially segregated and with higher black shares tend to have a larger decrease in rents (Table 3), a larger increase in black share (Table 4) and a larger increase in low-income share (Table 5). Interestingly, housing units located in tracts in MSAs with a relatively high Income Segregation Index appear to experience (a) smaller increases in the low-income share of existing housing units and (b) a smaller decrease in rents after new low-income housing units are built in that tract.

Columns 4 and 8 shows that interaction effects play an important role in explaining differences in impacts across tracts within MSAs. For instance, column 4 of table 3 shows a negative and significant coefficient on the MSA racial dissimilarity main effect (-0.074), a positive and significant coefficient on the interaction with tract black share (0.218), and a negative and significant coefficient on the interaction with tract poverty (-0.261). A similar pattern occurs for MSA black share and those same interactions. These patterns of coefficients suggest that negative rent impacts are especially pronounced in the poorer, non-black tracts in highly segregated MSAs with large MSA black shares.

6.3 Low-income housing and child human capital

In recent work, Chetty, Friedman, Hendren, Jones, and Porter (2018) document that the children of parents with identical incomes vary substantially in their adult incomes based on the Census tract of residence during childhood, suggesting that neighborhoods are an important input to human capital formation. The Chetty, Friedman, Hendren, Jones, and Porter (2018) estimates of the impact of tract on adult earnings, called tract "value added" throughout the rest of our paper, is publicly available as the *Opportunity Atlas.* The research of Chetty, Friedman, Hendren, Jones, and Porter (2018) has spurred a renewed interest in housing policies that encourage low-income households to reside in high-income-producing tracts as a means of promoting intergenerational income mobility.

We use simulations of our model, together with the tract-level Opportunity Atlas value-added measures, to quantify the impact of additional low-income housing units placed in any given tract on both the adult earnings of the children of the households who occupy those units as well as the children of other households in the tract and elsewhere in the MSA. The exercise involves three parts. First, we compute the new steady state of the model to forecast the demographic changes that are likely to occur in every tract of an MSA after low-income housing units are added to a specific tract in that MSA. Second, for all tracts in the MSA we adjust the Opportunity Atlas tract-level value-added estimates to account for all of the tract-level changes in demographic composition. Finally, we compute the aggregate tract-level value-added for all children in the metro area before and after the low-income housing units are built and report the change in this aggregate statistic.

We adjust the reported Opportunity Atlas estimates because the analysis in sections 6.1 and 6.2 suggests that some low-income housing developments can yield non-negligible steady-state changes to the demographic mix of neighborhoods.[26] We assume that the Opportunity

[26] A natural question whether the Opportunity Atlas neighborhood value added measures, estimated with data from an earlier time period, will accurately forecast how a neighborhood would contribute to child human capital formation after a large intervention and any resulting demographic changes. Chetty, Friedman, Hendren, Jones, and Porter (2018) show that the effects of tract on income are stable for different cohorts of children born from 1978-1989. This suggess estimates of neighborhood effects from an earlier time period would have provided a useful forecast of neighborhood effects in a later time period (in the historical policy

Atlas neighborhood value-added measure can be written as the sum of several additively seperable components,

$$VA_j = \varphi_{m(j)} + \xi_j + g\Big(B_j, L_j; \gamma_g\Big) \tag{16}$$

where $\varphi_{m(j)}$ is an MSA component common to all neighborhoods in the same MSA ($m(j)$ is the MSA containing j), ξ_j is an exogenous tract level component that is invariant to changes in neighborhood composition, and $g\Big(B_j, L_j; \gamma_g\Big)$ is an endogenous component that is a function of the location's demographic mix as in Diamond (2016). The term γ_g is a vector of parameters that characterizes the impact of demographics on a tract's value added.

We consider two cases for g as a means of providing bounds on what we can reasonably expect. First, we set $g = 0$. This assumes that neighborhood effects are entirely unaffected by tract demographic change. In the second case, we estimate the parameters of equation (16) by regressing the reported tract value-added on MSA fixed effects and a cubic spline in B_j and L_j (parameterized by γ_g). This approach overstates the impact of black- and low-income peers on value added if households sort on neighborhood characteristics that are correlated with value added, but not directly causal to value added.

The metro-level aggregate change in value added can come from three sources, which we individually track. First, we track the change in vaue added for the types of households that move into the low-income units. As with the simulations in sections 6.1 and 6.2, these households are assumed to be moving from outside the metro area.[27] Additionally, for each household, we assume that if the household had not moved into the new low-income units, that household would have received the average value-added for their type in the MSA with the new units. Second, we track the change in value added for the existing residents of the MSA that move in response to the policy. This includes all residents that move to a different tract, not just residents of the tract with the additional low-income units. Third, we track the change in value-added received by residents that do not move. This value added can change if neighborhood composition changes, depending on how the g function of equation (16) is specified.

Table 6 summarizes the results of simulations measuring the impact on the expected annual adult income of the MSA children when 100 low income units are added to a single tract. We repeat the experiment once for each tract in our sample treated as the target tract. Columns (1)-(3) report results when tract-value added is not affected by neighborhood

environment).

[27] Also as before, the types of households living in the newly built low-income housing units are all assumed to earn income in the bottom income tercile; and, the black share of these households is equal to the black share of households in the bottom income tercile in that MSA.

composition and $g = 0$. Column (1) reports the average of the aggregate MSA-total impact on *annual* adult earnings of \$11,888 with a standard deviation of the aggregate across the experiments (where the receiving tract differs) of \$232,157. The column shows that, on average, we find small impacts for both the occupants of the new low-income units (\$29,773) and for prior residents as the result of relocating (-\$17,885).

That said, the aggregate impact of adding low income units is positive for 50% of tracts. Columns 2 and 3 provide similar calculations separately when low-income housing is added to tracts with aggregate negative effects (column 2) and aggregate positive effects (column 3). The lower portion of these columns also shows characteristics of the tracts in those samples. The tracts with positive impacts have lower minority shares, lower poverty and higher median incomes, and lower rental shares. What these columns illustrate is that the low-income housing must be put in the right place to have a positive impact on value added, as the average annual aggregate impact for negative tracts is -\$169 thousand while it is more than \$191 thousand for the positive tracts. The third column also shows that, on average, the households that move into the new low-income housing units experience large gains; *on average*, the impact of the change on the households that relocate is modestly negative although the standard deviation is quite high at \$96 thousand.

Columns (4)-(6) report results when the g function of equation (16) is estimated as described and neighborhood composition directly affects value added. Figure 10 shows a scatterplot of the (steady state) before and after value-added measures for the specific tract receiving the 100 low-income units in each simulation. If a dot lies below the 45 degree line, it means that the steady-state value-added declines after the low-income housing is built. The figure shows that adding low income units tends to reduce tract value added and that larger declines occur in tracts with high value added as according to the Opportunity Atlas. Table 7 reports the results of a regression of post-development value added on Opportunity Atlas value added and obervable neighborhood traits. Conditional on baseline value added, larger declines occur in tracts with a higher white share, a lower rental share, and tracts in MSAs with more racial segregation.

After accounting for these endogenous neighborhood changes (column 4), the average aggregate MSA-total impact on annual adult earnings falls -\$69,052 with a standard deviation of \$262,919.[28] The impact on existing residents via neighborhood change is on average -\$40,470, and the impact for the occupants of new low income units falls to -\$10,697 on average. In this scenario the total aggregate impact of adding low income units remains positive for 37% of tracts. A comparison of columns 5 and 6 shows that, similar to the experiments where value-added is unaffected by demographic mix, tracts where adding low-

[28] As before, the stadard deviation is measured across experiments.

income housing generates a positive aggregate impact on children's later earnings tend to be more affluent and have higher Opportunity Atlas value-added measures. Also as before, on average (but with a large standard deviation) the gains to value added accrue to the new residents of the additional low-income units. On average, the residents that relocate and the residents that choose not to move experience modest declines to value added.

7 Conclusions

We estimate the structural parameters of a location choice model using panel data on the location decisions of 5% of the US population and a regression discontinuity approach exploiting the eligibility of a tract for low-income housing tax credits for developers. For a randomly selected set of 839 tracts from every MSA with at least 100 tracts and no more than 250 tracts, we compute the steady state of the model when a relatively large number of low-income housing units are built in a given tract (policy 1) and when a relatively small number of low-income housing units are built in a cluster of geographically proximate tracts (policy 2). When measured at the median tract, rents fall and, for the existing units, the number of low-income residents rises whereas the share of black residents stays approximately flat. When the low-income units are placed in exactly one tract, the variance of potential outcomes is quite high and the specific tract receiving the units might matter considerably for outcomes. In contrast, when the new units are placed in a cluster of nearby tracts, the specific tract that is in the middle of the cluster does not seem to matter and the range of outcomes is quite small around the median outcome. If the goal of low-income housing policy is to increase the share of low-income households in a target tract, then a policy that distributes a small amount of low-income housing in a number of tracts proximate to the target tract is likely to achieve that outcome with little variance.

References

Baum-Snow, N., and J. Marion (2009): "The effects of low income housing tax credit developments on neighborhoods," *Journal of Public Economics*, 93(5-6), 654–666. 2, 5, 6

Bayer, P., F. Ferreira, and R. McMillan (2007): "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *Journal of Political Economy*, 115(4), 588–638. 2, 12

Bayer, P., R. McMillan, A. Murphy, and C. Timmins (2015): "A Dynamic Model of Demand for Houses and Neighborhoods," Duke University Working Paper. 7, 8

Bishop, K. C. (2012): "A Dynamic Model of Location Choice and Hedonic Valuation," Working Paper, Washington University in St. Louis. 9

Bishop, K. C., and A. D. Murphy (2011): "Estimating the Willingness to Pay to Avoid Violent Crime: A Dynamic Approach," *American Economic Review*, 101(3), 625–629. 8

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," . 10

Boustan, L. P. (2010): "Was Postwar Suburbanization White Flight? Evidence from the Black Migration*," *The Quarterly Journal of Economics*, 125(1), 417–443. 12

Card, D. (2001): "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 19(1), 22–64. 12

Chetty, R., J. N. Friedman, N. Hendren, M. R. Jones, and S. R. Porter (2018): "The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility," . 1, 3, 22, 33, 43

Davis, M. A., J. Gregory, D. A. Hartley, and K. T. K. Tan (2017): "Neighborhood Choices, Neighborhood Effects and Housing Vouchers," Unpublished Working Paper. 2, 7, 9, 10, 12

Diamond, R. (2016): "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, 106(3), 479–524. 23

Diamond, R., and T. McQuade (2017): "Who Wants Affordable Housing in their Backyard? An Equilibrium Analysis of Low Income Property Development," Forthcoming, Journal of Political Economy. 2, 3

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. 9

Kennan, J., and J. R. Walker (2011): "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 79(1), 211–251. 7, 8

Table 1: Balance of pre-Determined Tract Characteristics at $\mathcal{X}_j = 0$

2000 Census Measure	Mean Value	Change at $\mathcal{X}_j = 0$	
Percent of units, rental occupied	41%	1.36 ppt	(0.72 ppt)
Percent of units, vacant	7%	0.22 ppt	(0.26 ppt)
Median monthly rent ($'s)	625	-10.86	(7.57)
Percent Black	14%	0.99 ppt	(1.03 ppt)
Percent Hispanic	11%	0.04 ppt	(0.88 ppt)
Average family income ($1,000's)	62	-3.64	(1.00)
Poverty rate	13%	0.06 ppt	(0.09 ppt)
Percent receiving public assistance	9%	-0.16 ppt	(0.22 ppt)

Table 2: Impact of $\mathcal{X}_j = 0$ (as compared to $\mathcal{X}_j < 0$) on the Probability that a Tract is Chosen, by Race, Age and Income of Mover

Demographic Subgroup			Neighborhoods Included in Estimation Sample			
				Neighborhood's Predominant Race in 2000		
Race	"Type" Income	Age	All Neighborhoods	Black	Hispanic	White/Other
Black			0.151 (0.06)**	0.173 (0.162)	-0.076 (0.097)	0.106 (0.062)*
Hispanic			0.104 (0.042)**	-0.171 (0.078)**	0.21 (0.107)**	0 (0.064)
White			0.022 (0.031)	-0.013 (0.026)	0.121 (0.044)***	-0.051 (0.057)
Other			-0.008 (0.077)	-0.081 (0.06)	0.097 (0.079)	-0.044 (0.165)
	Low Income		0.063 (0.028)**	-0.049 (0.056)	0.135 (0.065)**	-0.004 (0.048)
	Non-low income		0.035 (0.028)	0.015 (0.032)	0.113 (0.041)***	-0.039 (0.053)
		< 35	0.067 (0.034)**	-0.016 (0.043)	0.133 (0.056)**	0.013 (0.065)
		35-44	0.047 (0.026)*	0.011 (0.038)	0.114 (0.048)**	-0.014 (0.05)
		45-54	0.043 (0.025)*	-0.022 (0.042)	0.14 (0.049)***	-0.008 (0.046)
		55-64	0.006 (0.028)	-0.018 (0.043)	0.051 (0.047)	-0.058 (0.054)
		65+	-0.04 (0.034)	0 (0.046)	0.103 (0.046)**	-0.256 (0.067)***
Black	Low Income		0.156 (0.064)**	0.121 (0.177)	-0.045 (0.101)	0.122 (0.064)*
Black	Non-low income		0.146 (0.06)**	0.222 (0.158)	-0.106 (0.101)	0.09 (0.066)
Hispanic	Low Income		0.105 (0.042)**	-0.166 (0.077)**	0.204 (0.107)*	0.014 (0.065)
Hispanic	Non-low income		0.096 (0.059)	-0.248 (0.131)*	0.296 (0.14)**	-0.214 (0.097)**
White	Low Income		0.019 (0.033)	-0.032 (0.028)	0.099 (0.052)*	-0.039 (0.061)
White	Non-low income		0.023 (0.031)	-0.009 (0.027)	0.125 (0.043)***	-0.053 (0.057)
Other	Low Income		-0.046 (0.084)	-0.092 (0.063)	0.103 (0.089)	-0.188 (0.18)
Other	Non-low income		0.069 (0.077)	-0.056 (0.088)	0.087 (0.081)	0.238 (0.165)
Black		< 35	0.116 (0.062)*	0.163 (0.166)	-0.085 (0.11)	0.047 (0.072)
Black		35-44	0.177 (0.062)***	0.276 (0.16)*	-0.079 (0.103)	0.098 (0.067)
Black		45-54	0.151 (0.07)**	-0.031 (0.195)	0.067 (0.117)	0.142 (0.07)**
Black		55-64	0.138 (0.079)*	0.048 (0.224)	-0.132 (0.129)	0.169 (0.083)**
Black		65+	0.149 (0.087)*	0.292 (0.263)	-0.298 (0.142)**	0.191 (0.089)**
Hispanic		< 35	0.124 (0.046)***	-0.164 (0.086)*	0.232 (0.113)**	0.025 (0.074)
Hispanic		35-44	0.117 (0.044)***	-0.18 (0.079)**	0.215 (0.106)**	0.013 (0.069)
Hispanic		45-54	0.072 (0.049)	-0.098 (0.093)	0.178 (0.124)	-0.006 (0.075)
Hispanic		55-64	0.046 (0.054)	-0.233 (0.112)**	0.121 (0.135)	0.017 (0.083)
Hispanic		65+	0.096 (0.059)	-0.248 (0.131)*	0.296 (0.14)**	-0.214 (0.097)**
White		< 35	0.056 (0.042)	-0.029 (0.034)	0.132 (0.057)**	0.003 (0.079)
White		35-44	0.029 (0.032)	-0.005 (0.028)	0.118 (0.047)**	-0.02 (0.06)
White		45-54	0.03 (0.03)	-0.013 (0.028)	0.152 (0.044)***	-0.023 (0.056)
White		55-64	-0.012 (0.034)	0.002 (0.027)	0.068 (0.042)	-0.113 (0.066)*
White		65+	-0.078 (0.04)*	-0.014 (0.031)	0.101 (0.039)**	-0.31 (0.079)***
Other		< 35	0.092 (0.097)	-0.108 (0.074)	0.173 (0.083)**	0.222 (0.209)
Other		35-44	-0.076 (0.091)	-0.159 (0.07)**	0.087 (0.097)	-0.225 (0.192)
Other		45-54	0.073 (0.09)	0.239 (0.083)***	0.163 (0.092)*	-0.069 (0.2)
Other		55-64	0.047 (0.085)	0.025 (0.158)	0.002 (0.115)	0.246 (0.176)
	Low Income	< 35	0.081 (0.034)**	-0.077 (0.051)	0.174 (0.074)**	0.033 (0.061)
	Low Income	35-44	0.067 (0.032)**	-0.051 (0.054)	0.15 (0.07)**	-0.015 (0.057)
	Low Income	45-54	0.049 (0.031)	-0.072 (0.078)	0.133 (0.069)*	-0.016 (0.049)
	Low Income	55-64	0.042 (0.037)	-0.032 (0.099)	0.05 (0.081)	-0.027 (0.055)
	Low Income	65+	0.028 (0.055)	0.177 (0.162)	-0.208 (0.079)***	-0.042 (0.072)
	Non-low income	< 35	0.064 (0.038)*	0.016 (0.045)	0.11 (0.052)**	0.003 (0.072)
	Non-low income	35-44	0.046 (0.029)	0.036 (0.037)	0.1 (0.045)**	-0.003 (0.055)
	Non-low income	45-54	0.047 (0.03)	0.01 (0.03)	0.157 (0.044)***	0.002 (0.056)
	Non-low income	55-64	-0.006 (0.033)	-0.012 (0.028)	0.062 (0.041)	-0.072 (0.063)
	Non-low income	65+	-0.053 (0.036)	-0.04 (0.035)	0.15 (0.048)***	-0.298 (0.072)***

* Denotes significance at 10%, ** at 5% and *** at 1%.

Table 3: Regressions of change in log rent on tract- and msa-level characteristics

	Policy 1 (Mean impact = -0.078)				Policy 2 (Mean impact = -0.010)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tract variables (year 2000):								
Tract black share	-0.041***		-0.007***	-0.141***	-0.007***		-0.001*	-0.016***
	(0.00)		(0.00)	(0.01)	(0.00)		(0.00)	(0.00)
Tract poverty rate	0.017***		-0.021***	0.047*	-0.001		-0.007***	0.001
	(0.00)		(0.00)	(0.03)	(0.00)		(0.00)	(0.01)
Tract median rent ($1,000s)	0.017***		0.009***	0.015	0.003***		0.002***	0.000
	(0.00)		(0.00)	(0.01)	(0.00)		(0.00)	(0.00)
MSA variables (year 2000):								
Racial dissimilarity index (RDI)		-0.084***	-0.083***	-0.074***		-0.010***	-0.010***	-0.008***
		(0.00)	(0.00)	(0.01)		(0.00)	(0.00)	(0.00)
RDI x Tr. black share				0.218***				0.025***
				(0.02)				(0.00)
RDI x Tr. poverty rate				-0.228***				-0.034***
				(0.04)				(0.01)
RDI x Tr. median rent ($1,000s)				0.013				0.003
				(0.02)				(0.00)
Income segregation index (ISE)		0.202***	0.185***	0.180***		0.028***	0.024***	0.007
		(0.01)	(0.02)	(0.05)		(0.00)	(0.00)	(0.01)
ISE x Tr. black share				-0.579***				-0.071***
				(0.09)				(0.02)
ISE x Tr. poverty rate				1.174***				0.199***
				(0.20)				(0.04)
ISE x Tr. median rent ($1,000s)				-0.146**				-0.002
				(0.06)				(0.01)
MSA Black share		-0.144***	-0.131***	-0.172***		-0.027***	-0.024***	-0.031***
		(0.00)	(0.01)	(0.02)		(0.00)	(0.00)	(0.01)
MSA Blk share x Tr. black share				0.296***				0.037***
				(0.02)				(0.01)
MSA Blk share x Tr. poverty rate				-0.261***				-0.037***
				(0.04)				(0.01)
MSA Blk share x Tr. med. rent ($1,000s)				0.035				0.007
				(0.03)				(0.01)
Constant	-0.083***	-0.026***	-0.027***	-0.030***	-0.010***	-0.003***	-0.003***	-0.002
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	8,991	8,991	8,991	8,991	8,997	8,997	8,997	8,997
R-squared	0.082	0.192	0.209	0.252	0.087	0.152	0.181	0.200

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Regressions of change in black share on tract- and msa-level characteristics

	Policy 1 (Mean impact = +0.005)				Policy 2 (Mean impact = +0.000)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tract variables (year 2000):	0.015***		0.004***	0.024***	0.002***		0.000***	0.002***
Tract black share	(0.00)		(0.00)	(0.01)	(0.00)		(0.00)	(0.00)
	-0.009***		-0.001	0.003	-0.001***		0.000	0.002
Tract poverty rate	(0.00)		(0.00)	(0.01)	(0.00)		(0.00)	(0.00)
	-0.005***		-0.006***	-0.005	-0.000***		-0.000***	0.000
Tract median rent ($1,000s)	(0.00)		(0.00)	(0.01)	(0.00)		(0.00)	(0.00)
MSA variables (year 2000):								
Racial dissimilarity index (RDI)		0.003**	0.002	-0.002		0.000	-0.000	0.000
		(0.00)	(0.00)	(0.01)		(0.00)	(0.00)	(0.00)
RDI x Tr. black share				-0.019*				-0.001
				(0.01)				(0.00)
RDI x Tr. poverty rate				0.020				-0.001
				(0.02)				(0.00)
RDI x Tr. median rent ($1,000s)				0.002				-0.001
				(0.01)				(0.00)
Income segregation index (ISE)		-0.032***	-0.018**	-0.067***		-0.004***	-0.003***	-0.004*
		(0.01)	(0.01)	(0.02)		(0.00)	(0.00)	(0.00)
ISE x Tr. black share				0.036				0.001
				(0.05)				(0.00)
ISE x Tr. poverty rate				-0.051				-0.010
				(0.09)				(0.01)
ISE x Tr. median rent ($1,000s)				0.082***				0.004
				(0.03)				(0.00)
MSA Black share		0.048***	0.043***	0.113***		0.005***	0.004***	0.009***
		(0.00)	(0.00)	(0.02)		(0.00)	(0.00)	(0.00)
MSA Blk share x Tr. black share				-0.059***				-0.004***
				(0.01)				(0.00)
MSA Blk share x Tr. poverty rate				-0.089***				-0.003
				(0.02)				(0.00)
MSA Blk share x Tr. med. rent ($1,000s)				-0.075***				-0.005**
				(0.02)				(0.00)
Constant	0.007***	-0.001	0.002**	0.000	0.000***	0.000	0.000**	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	8,991	8,991	8,991	8,991	8,997	8,997	8,997	8,997
R-squared	0.054	0.108	0.117	0.146	0.058	0.108	0.118	0.128

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The change in black share refers only to the changes in thes share of the existing units.

Table 5: Regressions of change in low-income share on tract- and msa-level characteristics

	Policy 1 (Mean impact = +0.018)				Policy 2 (Mean impact = +0.001)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tract variables (year 2000):								
Tract black share	0.026***		-0.000	0.065***	0.003***		0.000*	0.004***
	(0.00)		(0.00)	(0.01)	(0.00)		(0.00)	(0.00)
Tract poverty rate	-0.013***		0.007*	-0.020	-0.001***		0.000	0.000
	(0.00)		(0.00)	(0.02)	(0.00)		(0.00)	(0.00)
Tract median rent ($1,000s)	-0.011***		-0.013***	-0.016*	-0.001***		-0.001***	-0.001*
	(0.00)		(0.00)	(0.01)	(0.00)		(0.00)	(0.00)
MSA variables (year 2000):								
Racial dissimilarity index (RDI)		0.019***	0.017***	0.001		0.000	0.000	-0.000
		(0.00)	(0.00)	(0.01)		(0.00)	(0.00)	(0.00)
RDI x Tr. black share				-0.092***				-0.006***
				(0.02)				(0.00)
RDI x Tr. poverty rate				0.116***				0.005**
				(0.03)				(0.00)
RDI x Tr. median rent ($1,000s)				0.011				0.000
				(0.02)				(0.00)
Income segregation index (ISE)		-0.056***	-0.029**	0.024		-0.007***	-0.005***	-0.003
		(0.01)	(0.01)	(0.04)		(0.00)	(0.00)	(0.00)
ISE x Tr. black share				0.175**				0.010
				(0.08)				(0.01)
ISE x Tr. poverty rate				-0.495***				-0.040***
				(0.15)				(0.01)
ISE x Tr. median rent ($1,000s)				-0.007				0.003
				(0.05)				(0.00)
MSA Black share		0.100***	0.097***	0.145***		0.009***	0.009***	0.011***
		(0.00)	(0.00)	(0.02)		(0.00)	(0.00)	(0.00)
MSA Blk share x Tr. black share				-0.133***				-0.006***
				(0.02)				(0.00)
MSA Blk share x Tr. poverty rate				0.006				0.003
				(0.04)				(0.00)
MSA Blk share x Tr. med. rent ($1,000s)				-0.037				-0.002
				(0.03)				(0.00)
Constant	0.021***	-0.002	0.003*	0.004	0.001***	0.000**	0.001***	0.001
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	8,991	8,991	8,991	8,991	8,997	8,997	8,997	8,997
R-squared	0.050	0.134	0.143	0.165	0.060	0.144	0.152	0.160

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The change in low income share refers only to the change in these share of the existing units.

Table 6: Aggregate impacts of low-income housing on children's adult earnings

	(1)	(2)	(3)	(4)	(5)	(6)
	Policy-Invariant Neighborhood Effects			Endogenous Neighborhood Effects		
	All tracts	Total Impact < 0	Total Impact > 0	All tracts	Total Impact < 0	Total Impact > 0
Aggregate impacts on annual adult income ($):						
Total impact	11,888	-169,216	190,657	-69,052	-219,674	186,375
	(232,157)	(122,654)	(167,017)	(262,919)	(165,662)	(189,874)
Impact for occupents of new low-income units	29,773	-144,844	202,140	-10,697	-144,185	215,673
	(223,978)	(124,903)	(156,378)	(231,391)	(151,739)	(154,265)
Impact from relocation of other households	-17,885	-24,371	-11,483	-17,885	-22,252	-10,480
	(71,119)	(27,476)	(96,053)	(71,119)	(27,476)	(97,294)
Impact from neighborhood change	0	0	0	-40,470	-53,237	-18,818
	(-)	(-)	(-)	(73,106)	(57,698)	(89,546)
Total impact > 0	0.50	1.00	0.00	0.37	1.00	0.00
Neighborhood demographics:						
2000 Census:						
Black share		0.28	0.06		0.22	0.08
Hispanic share		0.10	0.05		0.09	0.05
Median household income		30,991	48,297		33,577	50,085
		(11,803)	(15,792)		(12,475)	(17,081)
Poverty rate		0.21	0.08		0.18	0.08
Share receiving public assistance		0.13	0.05		0.11	0.05
Rental occupied share		0.39	0.23		0.36	0.23
Owner occupied share		0.51	0.70		0.55	0.70
Vacant share		0.10	0.07		0.09	0.07
Median rent		0.49	0.61		0.51	0.63
Opportunity Atlas:						
Inc. percentile | parents inc. 25th percentile		0.36	0.45		0.37	0.46
Observations	8,942	4,442	4,500	8,942	5,625	3,317

Table 7: Regression of neighborhood effects on children's eventual adult incomes on tract- and msa-level variables

	(1)
Opp Atlas:	
Inc. percentile | parents inc. 25th percentile	0.901***
	(0.01)
Tract variables (2000 Census)	
White share	-0.004***
	(0.00)
Poverty rate	-0.001
	(0.00)
Rental occupied share	-0.009***
	(0.00)
Vacant share	0.005*
	(0.00)
Median rent	-0.001
	(0.00)
MSA variables (2000 Census)	
MSA racial segregation index	-0.013***
	(0.00)
MSA income segregation index	0.028***
	(0.01)
Constant	0.040***
	(0.00)
Observations	8,942
R-squared	0.886

Standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Note: There is one observation per Census tract. The dependent variable is the average nationwide percentile of the adult income of the tract's children conditional on having parents with income at the 25th nationwide percentile (Chetty, Friedman, Hendren, Jones, and Porter, 2018), adjusted for the simulated impact of placing 100 low income units in the tract on the tract's black share and low income share. The explanatory variables are pre-determined MSA and tract characteristics measured in the 2000 Census.

Figure 1: QCT Status by Tract Poverty/Income Running Variable

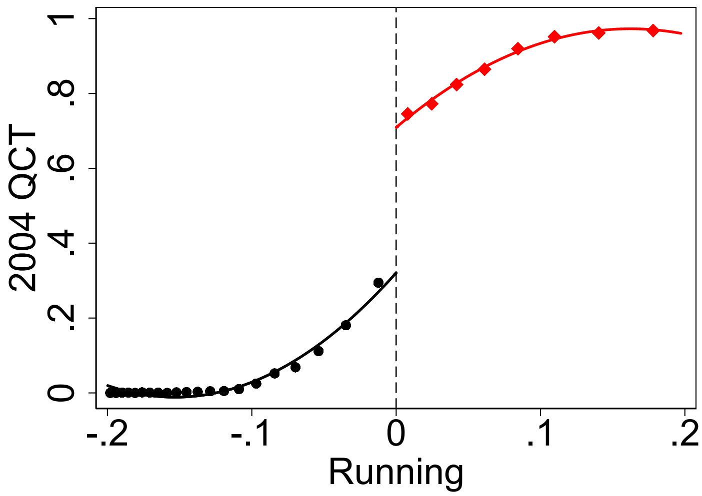
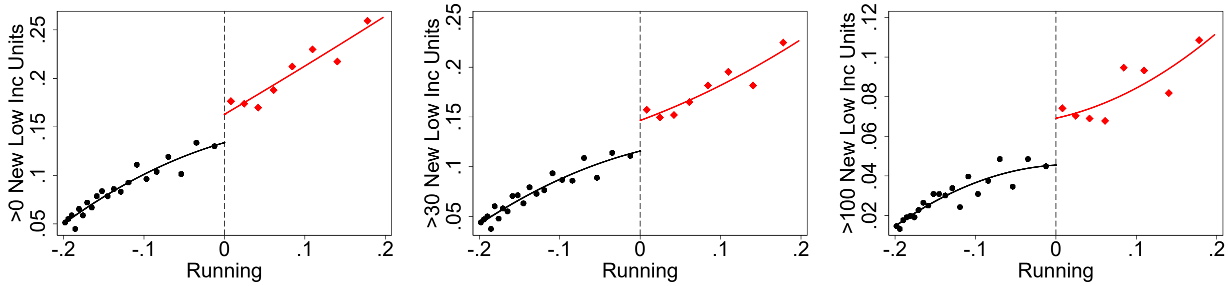


Figure 2: Post-2004 Low-Income Housing Development by Neighborhood Poverty and Median Income



(a) Any New Low Income Units (b) $>$ 30 New Low Income Units (c) $>$ 100 New Low Income Units

Figure 3: Density of Neighborhoods by Median Income (Index) and Poverty Rate

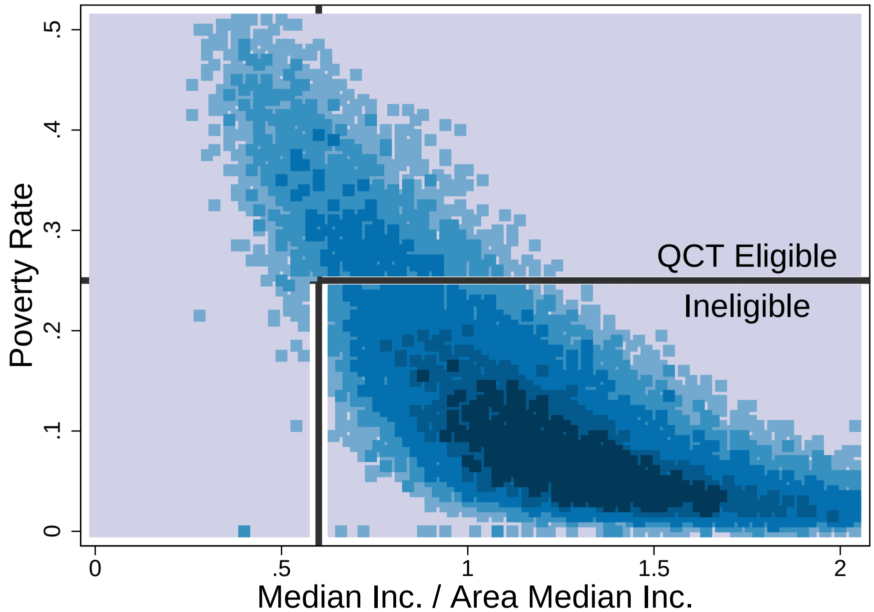
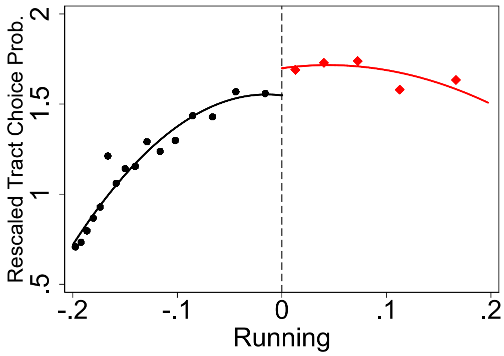
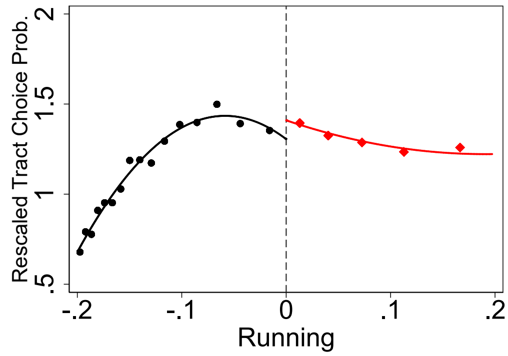


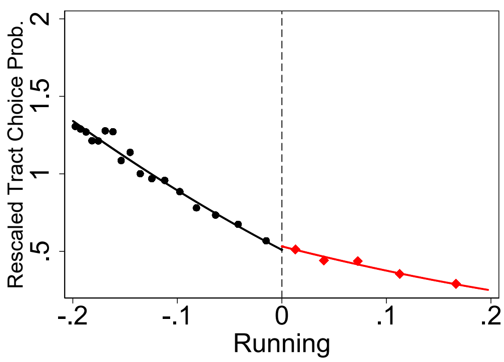
Figure 4: Rescaled Tract Choice Probabilities by $\mathcal{X}_j$



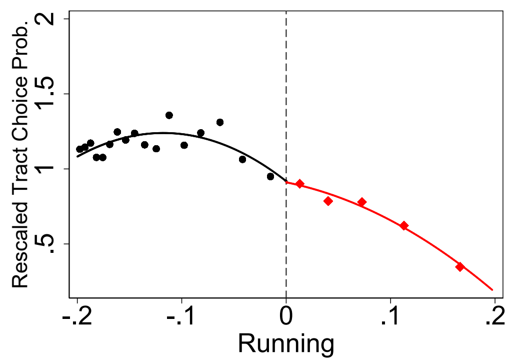
(a) Black HHs



(b) Low-Income, Hisp. HHs



(c) White HHs



(d) Other-race HHs

Figure 5: Household Type's Indirect Utility from QCT Eligibility: Distrubtion of t-Statistics

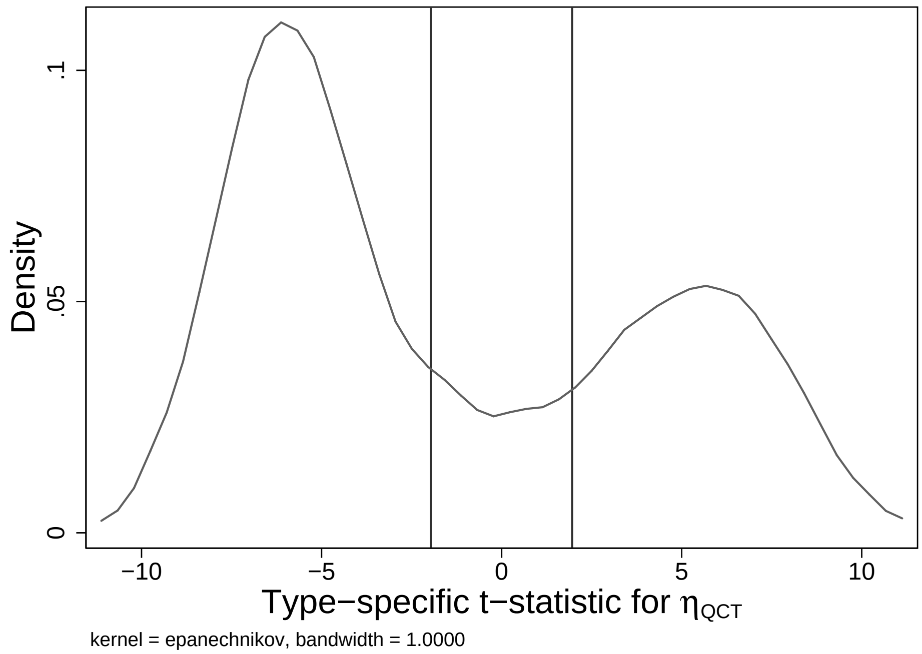
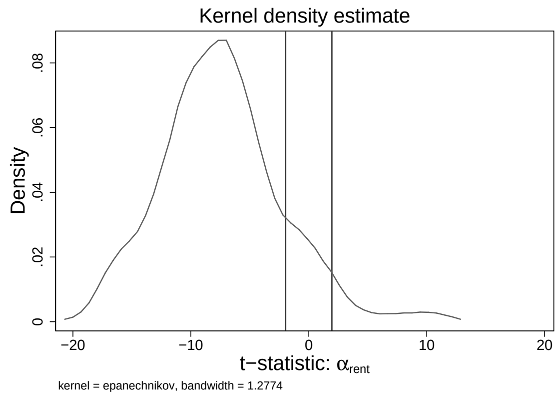
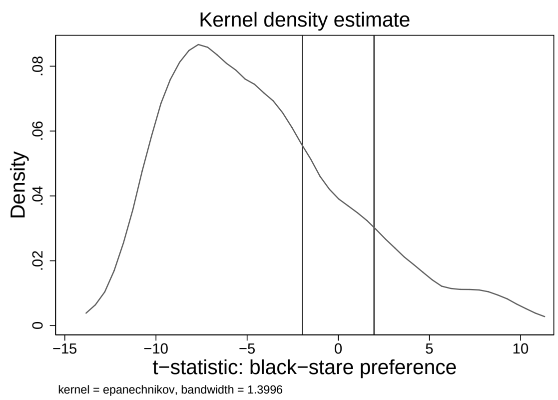


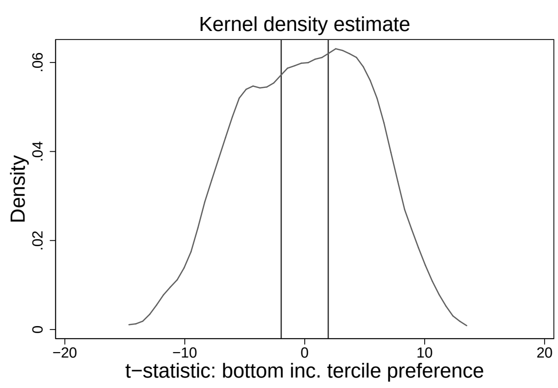
Figure 6: Household Type's Indirect Utility from Endogenous Neighborhood Traits: Distrubtion of t-Statistics



(a) log(rent)

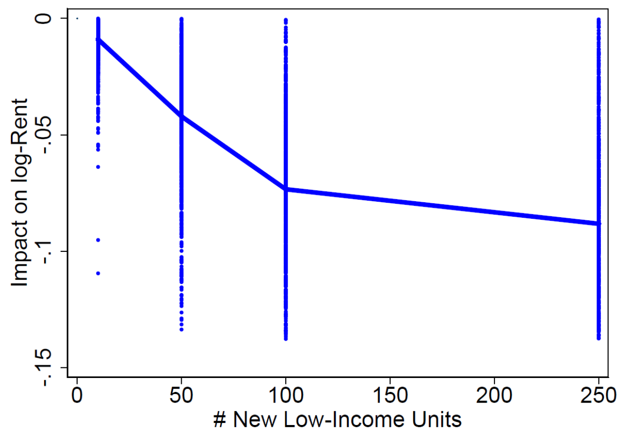


(b) Black share

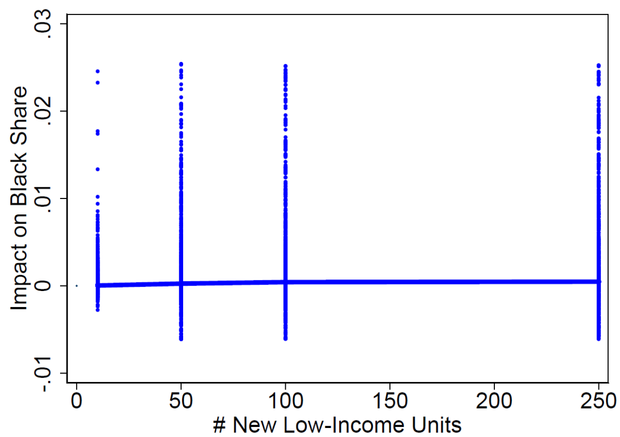


(c) Share in 1st income tercile

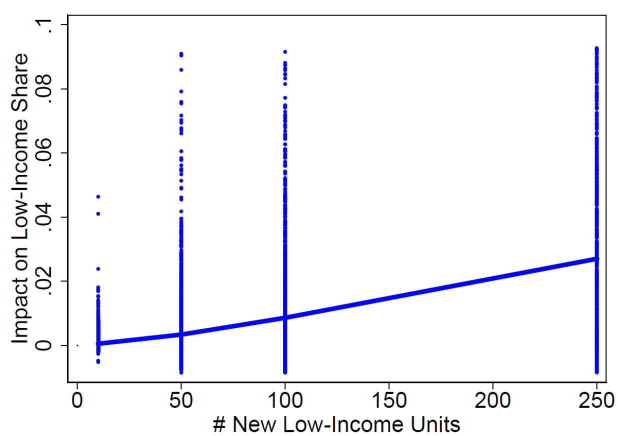
Figure 7: Impact of 10, 50, 100 and 250 Low-Income Housing Units Built in One Tract



(a) Δ Log Rent



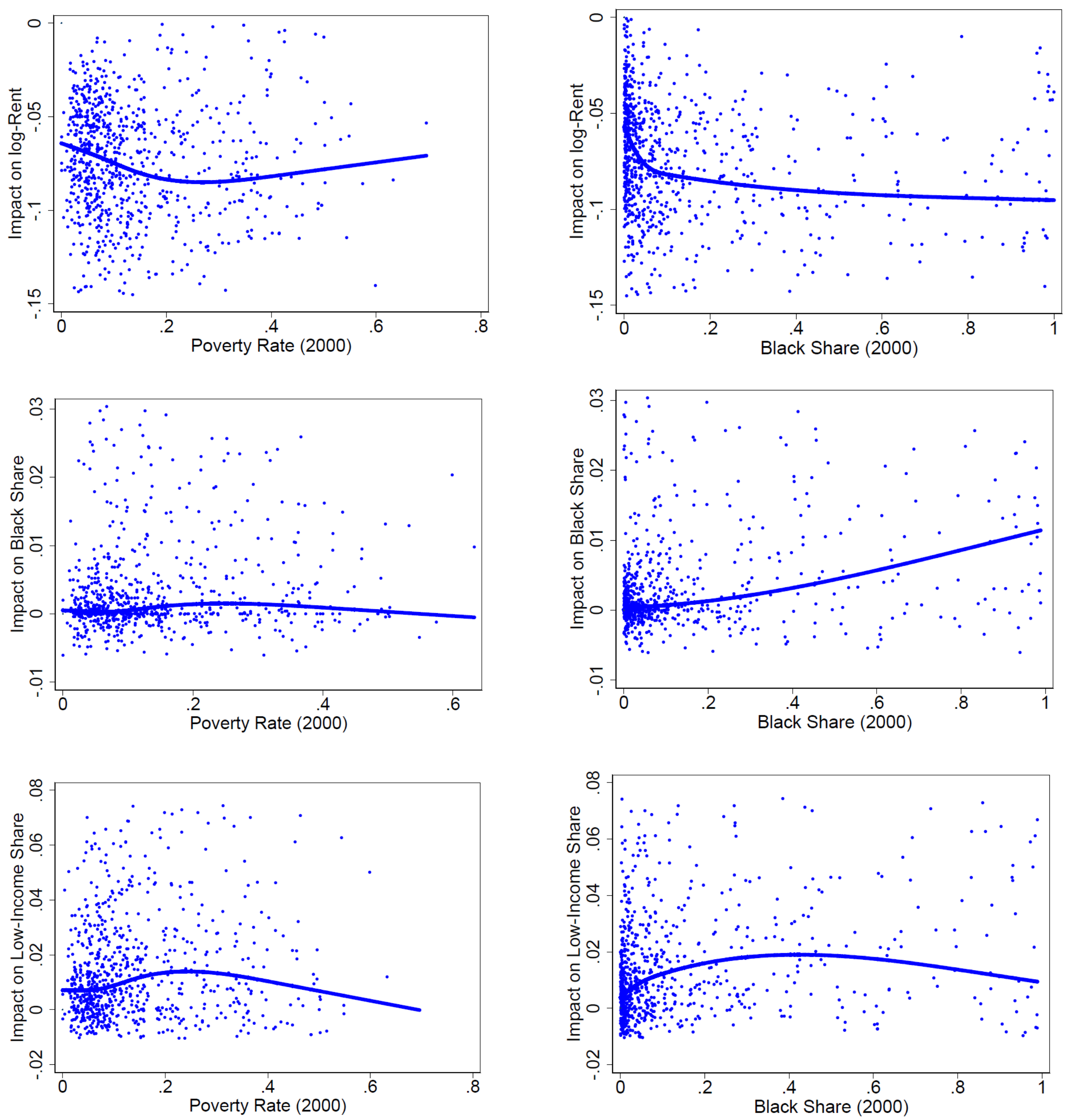
(b) Δ Black Share



(c) Δ Low-Income Share

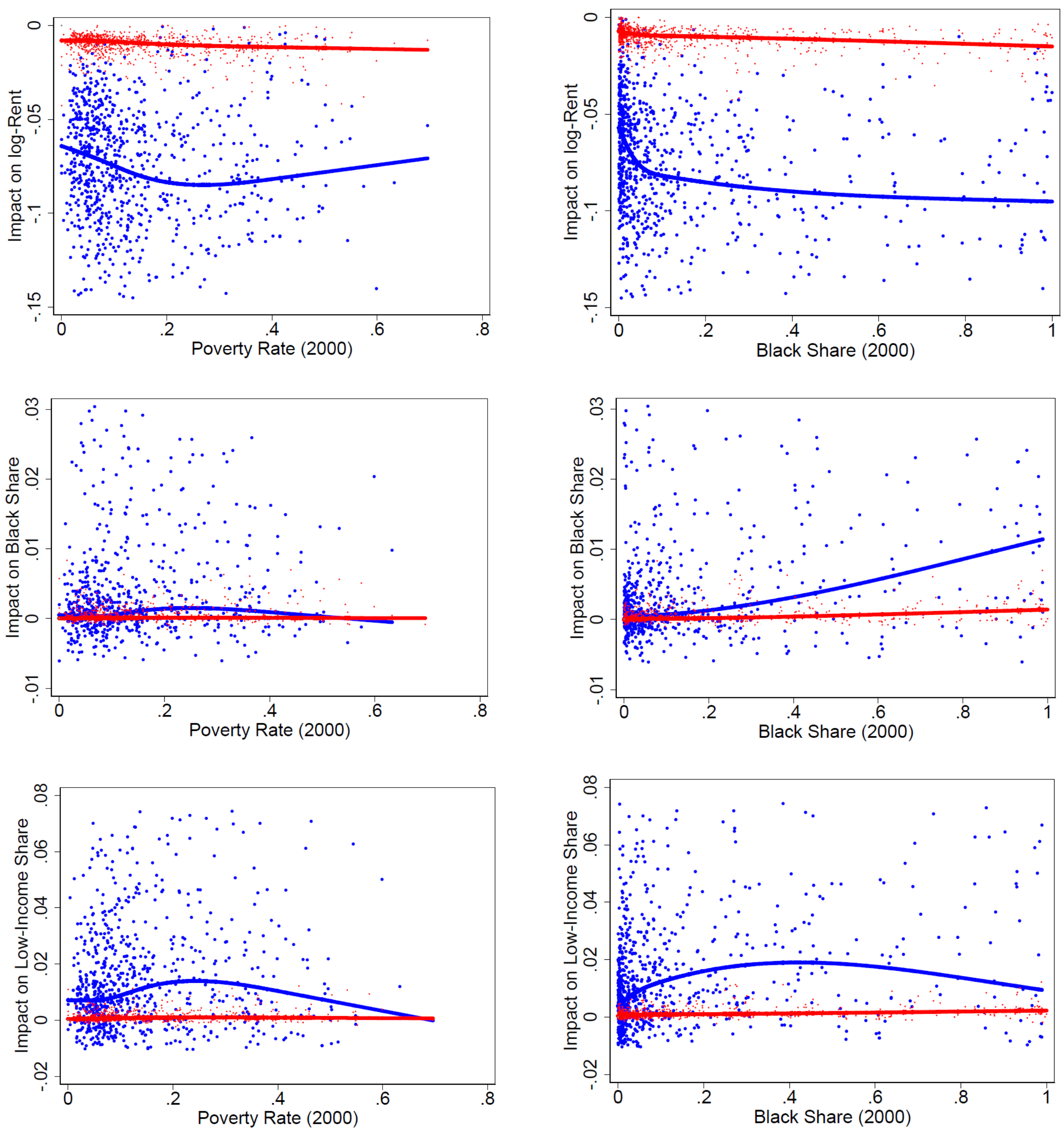
Note: Δ Black Share and Δ Low Income Share refer only to the changes in these shares of the existing units.

Figure 8: Impact of 100 Low-Income Housing Units Built in One Tract by Tract-Level Poverty Rate (Left) and Black Share (Right)



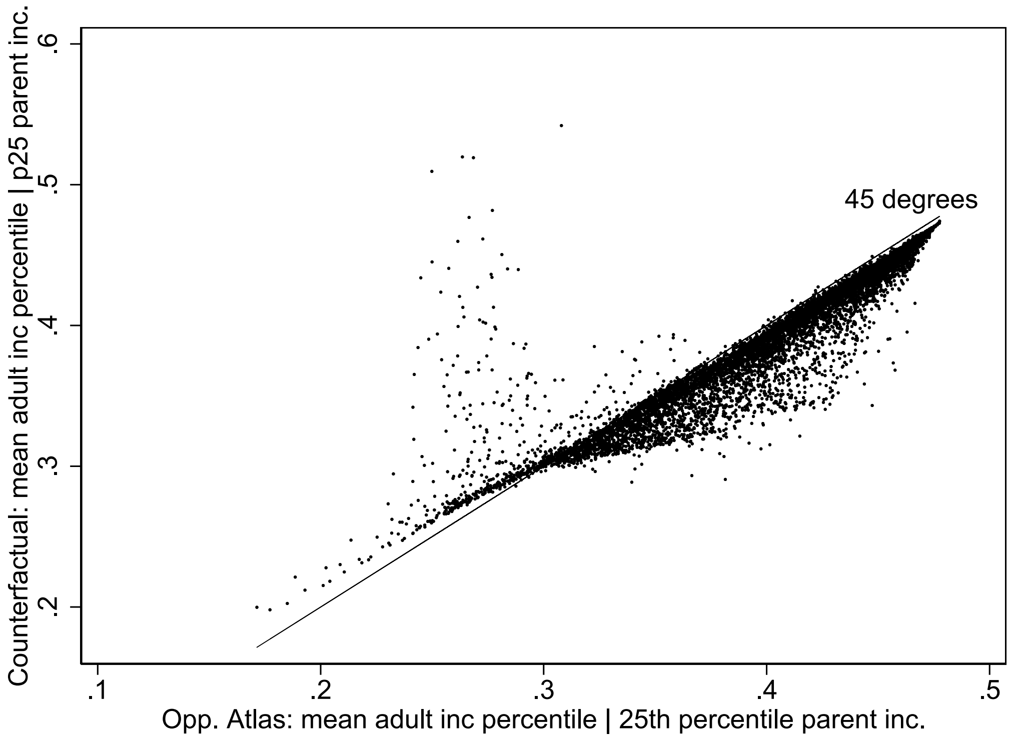
Note: Δ Black Share and Δ Low Income Share refer only to the changes in these shares of the existing units.

Figure 9: Impact of 10 Low-Income Housing Units Built in Target Tract and Surrounding Tracts by Tract-Level Poverty Rate (Left) and Black Share (Right)



Note: Δ Black Share and Δ Low Income Share refer only to the changes in these shares of the existing units.

Figure 10: The average adult income percentile of tract children before and after adding 100 low income units



Note: Each dot is one Census tract. The x-axis shows the average nationwide percentile of the adult income of the tract's children conditional on having parents with income at the 25th nationwide percentile (Chetty, Friedman, Hendren, Jones, and Porter, 2018). The y-axis shows the predicted value of the same measure after placing 100 low income units in the tract.