Neighborhood Choices and Neighborhood Effects*

Morris A. Davis
Rutgers University
morris.a.davis@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

Kegon T. K. Tan
University of Wisconsin - Madison
ttan8@wisc.edu

October 12, 2015

[**Preliminary and Incomplete: Do Not Cite**]

Abstract

We investigate how households choose where to live and how neighborhoods affect the ability of children. We use detailed panel data to estimate a dynamic model of neighborhood choice at the Census tract level for renters in Los Angeles county. We then use different panel data for Los Angeles to estimate tract-level "neighborhood effects," defined as the impact of neighborhoods on child cognitive ability. We conclude by simulating a Moving-to-Opportunity type experiment with our model, in which people residing in high poverty neighborhoods are given a rental voucher to move to a low-poverty Census tract. Child ability does not improve in these simulations, as households receiving vouchers tend to move to the least expensive eligible neighborhoods with the lowest neighborhood effects. If these households had chosen a Census tract randomly among the eligible set, child ability would have improved significantly.

JEL Classification Numbers: TBD
Keywords: Neighborhood Choice, Neighborhood Effects, Moving to Opportunity

*We thank Dionissi Aliprantis and numerous seminar participants for helpful comments. 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

In this paper we investigate how households optimally choose a neighborhood in which to live and how neighborhoods affect the cognitive ability of children. These topics have been studied individually before, but our approach is different and our data are new. We show that neighborhoods vary in their impact on child ability, and parents differ in willingness to pay to move to neighborhoods that may significantly improve child ability. We conclude the paper by showing that households differ in the sensitivity of their optimal neighborhood choice to rental prices and this is key to understanding the empirical results of the Moving-to-Opportunity experiment.

Our paper has three main sections, and the first two reflect contributions to heretofore distinct literatures. In our first section, we specify and estimate a dynamic model of optimal location choice using detailed micro panel data, in the spirit of Kennan and Walker (2011) and Bayer, McMillan, Murphy, and Timmins (2015). We estimate the model using panel data from the Federal Reserve Bank of New York (FRBNY) Consumer Credit Panel / Equifax. This is a 5% random sample of U.S. adults with an active credit file and any individuals residing in the same household. To our knowledge we are the first to use these data to estimate a location choice model. We restrict our sample to renters residing in Los Angeles County. We study renters to mitigate the influence of availability of credit on location choice, and we focus on Los Angeles County to match our results with estimates of the impact of neighborhoods on child ability, discussed next. Our estimation sample from the FRBNY Consumer Credit Panel / Equifax data consists of more than 1.75 million person-year observations. This huge sample allows us to estimate a full vector of model parameters for many discrete "types" of people. Our use of many types in estimation minimizes the role of unobservable shocks in explaining differences in optimal location choices. We show for many types of household, utility varies greatly across Census tracts; and, for many Census tracts, the utility of living in the tract varies widely across types.

In our second section, we estimate the impact of neighborhoods, in our case specific Census tracts in Los Angeles county, on the cognitive ability of children. There is a large literature in the social sciences studying these "neighborhood effects" on child ability, adolescent behavior, health, labor earnings, and other individual level outcomes. Empirical studies using observational data often find strong associations between neighborhood quality, broadly defined, and positive individual-level outcomes: See Leventhal and Brooks-Gunn (2000) and Durlauf (2004) for recent surveys. While these studies typically attempt to account for selection issues,[1] the fact that individuals endogenously sort into neighborhoods

[1]For example, Cutler and Glaeser (1997) study the impact of segregation on outcomes of African-

leaves open the possibility of non-causal explanations for these patterns.[2]

We make two contributions to this literature. First, we use a new longitudinal dataset in estimation, the Los Angeles Family and Neighborhood Survey (LA FANS). The LA FANS data allow for substantially richer controls than are typically available in observational studies of neighborhood effects. Second, we estimate the impact of neighborhoods on child ability using a "value-added approach", in which changes in student ability over time, as measured by changes in math and verbal test scores, are regressed on neighborhood fixed effects and a set of individual-level controls including, most importantly, lagged child test scores. The value-added approach has been applied widely in assessing teacher quality (for example Kane and Staiger (2008) and Chetty, Friedman, and Rockoff (2014)) but has not yet been used in the neighborhood effects literature.

The key advantage of the value-added approach for our application is that the method recovers estimates of the effect of specific neighborhoods on child ability, as compared to the average effect of neighborhoods associated with particular observable characteristics such as average income level and racial composition (the typical approach in the neighborhood-effects literature). We estimate economically important variation in neighborhood value-added across Census tracts in Los Angeles County: Variation in the neighborhood value-added that children are exposed to between waves 1 and 2 of the LA FANS data explains about 5% of the cross-sectional variance in child ability. In support of a causal, as opposed to selection-driven, interpretation of our neighborhood value-added estimates, we show that after we have controlled for children's lagged test scores and demographics, controlling additionally for variables such as parental ability, parental demographics, and household income and assets, which are strongly predictive of child ability in simple cross-sectional regressions, add very little in explanatory power for changes over time in child test scores.

In the third section, we reconcile the contradictory conclusions of the "neighborhood effects" literature and the papers that study the Moving-to-Opportunity (MTO) experiment. The Moving to Opportunity experiment was a randomized control trial beginning in the 1990s that randomly assigned a group of households eligible to live in low income housing projects in five U.S. cities to three different groups; (i) a treatment group that received a Section 8 housing voucher that in the first year could be applied only in Census tracts with a poverty rate under 10% and could be applied unconditionally thereafter, (ii) a second treatment group that received a comparable Section 8 housing voucher with no location requirement attached,

Americans using topographical features of cities as instruments for location choice and Aaronson (1998) measures neighborhood effects by studying outcomes of siblings at least three years apart in age after a move.

[2]See Aaronson (1998) for examples of instruments used by other researchers in this field and their potential limitations.

and (iii) a control group that received no voucher. Summarizing the medium to long term impacts of MTO, Sanbonmatsu, Kling, Duncan, and Brooks-Gunn (2006), Kling, Liebman, and Katz (2007) and Ludwig, Duncan, Gennetian, Katz, Kessler, Kling, and Sanbonmatsu (2013) show that on average the MTO treatment successfully reduced exposure to crime and poverty and improved the mental health of female children, but failed to improve child ability, educational attainment, or physical health.[3]

Many view the results from MTO as evidence against the hypothesis that neighborhoods can have large effects on a child's development of skills and educational attainment. Our view is that the results are open to multiple interpretations. One possibility is that, indeed, the findings of large neighborhood effects from earlier observational studies are driven entirely by selection of families into neighborhoods, and that true but unobserved neighborhood effects are small or non-existant. A second interpretation is that families receiving a voucher in the MTO experiment chose their neighborhoods endogenously after considering how that voucher changed the relative price of eligible neighborhoods. Thus, depending on how neighborhoods were chosen by MTO participants after receiving a voucher, the intent-to-treat effect of the MTO subsidy offer may have differed substantially from the average treatment effect of lower poverty neighborhoods on outcomes: See Aliprantis (2015), Clampet-Lundquist and Massey (2008) and Pinto (2014) for related discussions.

In the spirit of Davis and Foster (2005) and Todd and Wolpin (2006), we run counterfactual simulations of our decision model to better understand why the MTO experiment did not improve child outcomes if neighborhood effects are in fact important.[4] To start, we replicate the environment created by the MTO experiment. We refer to this simulation as "MTO-A." To implement MTO-A, we require an estimate, by type, of how the utility of each neighborhood in Los Angeles would change given the change in rental prices induced by the voucher. We extract an estimate of the rental-price sensitivity of each of our types of households using the instrumental variables approach of Bayer, Ferreira, and McMillan (2007).

In MTO-A, we compute the optimal neighborhood choices of households that begin the sample in neighborhoods with public housing developments but are offered a rental voucher valid for use in low-poverty-rate neighborhoods. Importantly, we show that households that use the voucher move to low-poverty neighborhoods with the lowest neighborhood value-added, on average. Since the chosen neighborhoods have low value-added, child test scores do not improve. So, why does the model predict this outcome? We find that the types of

[3]Recent work by Chetty, Hendren, and Katz (2015) argues that MTO positively affected adult wages.

[4]Galiani, Murphy, and Pantano (2015) estimate a structural model of location choice directly using MTO data, and run counterfactual experiments with their estimated model, but do not study the impact of MTO on child well-being.

households likely to receive an MTO voucher are very sensitive to rental prices; and the highest-value-added neighborhoods in low-poverty tracts are the most expensive. To prove that selection on rental prices is key, we perform another simulation we call "MTO-B" in which we randomly assign MTO-eligible households to neighborhoods with similar poverty rates as those chosen in the MTO-A. Under this simulation, we predict that child test scores would have significantly improved. In other words, our counterfactual simulations suggest that parents use the MTO voucher to move to low-poverty neighborhoods, but they choose relatively cheap and low value-added neighborhoods after receiving the voucher and child outcomes do not improve.

2 Location Choice Model

We consider 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. 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 family can choose 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

$$V(\ell \mid j, \epsilon_\ell) \;=\; u(\ell \mid j, \epsilon_\ell) + \beta EV(\ell) \tag{1}$$

In the above equation $EV(\ell)$ is the expected future value of having chosen to live in ℓ today. 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

$$u(\ell \mid j, \epsilon_\ell) \;=\; \delta_\ell \;-\; \kappa \cdot 1_{\ell \neq j} \;+\; \epsilon_\ell \tag{2}$$

δ_ℓ is the flow utility the household receives this period from living in neighborhood ℓ, net of rents and other costs; κ is the sum of all costs (utility and financial) a household must pay when it moves to a different neighborhood i.e. when $\ell \neq j$; 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 δ_ℓ and κ 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

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

$EV\left(j\right)$ 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 and allow κ to vary with distance. Bishop and Murphy (2011) and Bayer, McMillan, Murphy, 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 and across households, 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 $EV\left(j\right)$ has the functional form

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

where ζ is equal to Euler's constant and

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

That is, the tilde symbol signifies that the shock ϵ_ℓ has been omitted. Additionally, it can be shown that the log of the probability location ℓ is chosen given a current location of j, call it $p(\ell \mid j)$, has the solution

$$p(\ell \mid j) \;=\; \widetilde{V}(\ell \mid j) \;-\; \log\left\{\sum_{\ell'=1}^{J} \exp\left[\widetilde{V}(\ell' \mid j)\right]\right\} \tag{6}$$

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

$$p(\ell \mid j) \;=\; \widetilde{V}(\ell \mid j) - \widetilde{V}(k \mid j) \;-\; \log\left\{\sum_{\ell'=1}^{J} \exp\left[\widetilde{V}(\ell' \mid j) - \widetilde{V}(k \mid j)\right]\right\} \tag{7}$$

One approach to estimating model parameters such as Rust (1987) is to solve for the value functions at a given set of parameters, apply equation (7) directly to generate a likelihood over the observed choice probabilities, and then search for the set of parameters that maximizes the likelihood. This approach is computationally intensive because it requires solving for the value functions at each step of the likelihood, which involves backwards recursions using equation (4). In cases such as ours, involving many parameters to be estimated, this approach is computationally infeasible.

Instead, we use the approach of Hotz and Miller (1993) and employed by Bishop (2012) in similar work to proceed. This approach does not require that we solve for the value functions. Note that equation (5) implies

$$\widetilde{V}(\ell \mid j) - \widetilde{V}(k \mid j) \;=\; \delta_\ell - \delta_k \;-\; \kappa\left[1_{\ell \neq j} - 1_{k \neq j}\right] \;+\; \beta\left[EV(\ell) - EV(k)\right] \tag{8}$$

But from equation (4),

$$EV(\ell) - EV(k) \;=\; \log\left\{\sum_{\ell'=1}^{J} \exp \widetilde{V}(\ell' \mid l)\right\} - \log\left\{\sum_{\ell'=1}^{J} \exp \widetilde{V}(\ell' \mid k)\right\} \tag{9}$$

Now note that equation (6) implies

$$p(k \mid \ell) \;=\; \widetilde{V}(k \mid \ell) \;-\; \log\left\{\sum_{\ell'=1}^{J} \exp\left[\widetilde{V}(\ell' \mid \ell)\right]\right\} \tag{10}$$

$$p(k \mid k) \;=\; \widetilde{V}(k \mid k) \;-\; \log\left\{\sum_{\ell'=1}^{K} \exp\left[\widetilde{V}(\ell' \mid k)\right]\right\} \tag{11}$$

and thus

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

is equal to

$$\begin{array}{rccc} & \widetilde{V}\left(k\mid\ell\right)-\widetilde{V}\left(k\mid k\right) & - & \left[p\left(k\mid\ell\right)-p\left(k\mid k\right)\right] \\ = & -\kappa\cdot 1_{\ell\neq k} & - & \left[p\left(k\mid\ell\right)-p\left(k\mid k\right)\right] \end{array} \tag{12}$$

The last line is quickly derived from equation (5). Therefore,

$$EV\left(\ell\right)-EV\left(k\right) \;=\; -\left[p\left(k\mid\ell\right)-p\left(k\mid k\right)+\kappa\cdot 1_{\ell\neq k}\right] \tag{13}$$

and equation (8) has the expression

$$\begin{aligned} & \widetilde{V}\left(\ell\mid j\right)-\widetilde{V}\left(k\mid j\right) \\ & =\delta_{\ell}-\delta_{k}\;-\;\kappa\left[1_{\ell\neq j}-1_{k\neq j}\right]\;-\;\beta\left[p\left(k\mid\ell\right)-p\left(k\mid k\right)+\kappa\cdot 1_{\ell\neq k}\right] \end{aligned} \tag{14}$$

Combined, equations (7) and (14) show 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.

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 an active credit file and any individuals residing in the same household as an individual from that initial 5% sample.[5] For years 1999 to the present, 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 or education, and it does not contain information on income or assets although it does include a credit score (specifically, 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

[5]The data include all individuals with 5 out of the 100 possible terminal 2-digit SSN combinations. While the leading SSN digits are based on the birth year/location, the terminal SSN digits are essentially randomly assigned.

year. Other authors have used the FRBNY Consumer Credit Panel / Equifax data to study the relationship of interest rates, house prices and credit (see Bhutta and Keys (2015) and (Brown, Stein, and Zafar, 2013)) and the impact of natural disasters on household finances (Gallagher and Hartley, 2014), but we are the first to use this data to estimate an optimal location-choice model.

We restrict our sample to individuals who, from 1999 through 2013, are never observed outside of Los Angeles county and who never hold a home mortgage, yielding 1,787,558 person-year observations. We study renters to mitigate any problems of changing credit conditions and availability of mortgages during the sample window; and we study Los Angeles in particular to link our estimates of utility to measures of neighborhood effects on child outcomes we estimate for Census tracts in Los Angeles (to be discussed later). We exclude from our estimation Census tracts with fewer than 150 rental units and tracts that are sparsely populated in the northern part of the county. The panel is not balanced, as some individuals' credit records first become active after 1999.

An advantage of the size of our data is that we can estimate a full set of model parameters for many "types" of people, where we define a type of person based on observable demographic and economic characteristics. This stands in contrast to previous studies of neighborhood choice such as Bayer, McMillan, Murphy, and Timmins (2015) where, due to lack of data, the authors restrict variation in model parameters across the population.

We stratify households into types using an 8-step stratifying procedure. We begin with the full sample, and subdivide the sample into smaller "cells" based on (in this order): the racial plurality of the 1999 Census block of residence (4 bins),[6] 5 age categories (cutoffs at 30, 45, 55, and 65), number of adults 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 20,000 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. When all said and done, this procedure yields 144 types of households.

Overall, there are 1,748 tracts in our estimation sample. If we were to estimate a separate value of δ for each tract and for each type, this would require us to estimate more than 250,000 parameters. For parsimony, for each type we specify that the utility of location j, δ_j, is a

[6] For individuals who enter the sample after 1999, we classify them based on the racial plurality of the block where they are first observed.

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

$$\delta_j = \sum_{k=1}^{K} a_k B_k \left(lat_j, lon_j\right) \tag{15}$$

The B_k are parameter-less basis functions. We use $K = 100$ basis functions for each type, such that with 144 types we estimate $(100 + 1) \times 144 = 14,544$ 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, $\jmath_{it}$ 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 $\theta_\tau = (a_1, a_2, \ldots, a_K, \kappa)$. The log likelihood of the sample is

$$\sum_{\tau} \sum_{i \in \tau} \sum_{t} \quad p\left(\ell_{it} \mid j_{it}; \theta_\tau\right) \tag{16}$$

$p(.)$ 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.

Due to our large number of types and tracts, it is impossible to report all parameter estimates. Instead, we summarize the estimates by examining the model's in-sample fit and compare the predictions of the model with all types included in estimation to the predictions the model estimated with fewer types. Table 1 reports actual annual cross-tract migration rates in our sample. About 8-1/2 percent of our sample moves to a different tract in each year, and that percentage falls from just above 11 percent for those under 30 to just above 3 percent for those aged 65 and above. The top panel of figure 1 compares the estimated model's predicted migration rates to these values. The model slightly overstates annual migration rates, but replicates the pattern of declining migration with age. The bottom panel of figure 1 plots annual average migration flows for each j-to-ℓ tract pair versus the model-predicted migration flows, inclusive of non-movers i.e. j-to-j flows. The scatter plot falls tightly along the 45-degree line.

Figure 2 compares the tract-to-tract flows predicted by our full model to the flows predicted by an alternative, similarly-estimated version of our model with just four types defined by the plurality of race/ethnicity (white, black, Hispanic, and other) of the Census tract of first residence. The top panel of figure 2 compares the two models' predicted non-migration shares across tracts. The two model's predictions along this dimension are closely, though not perfectly, aligned. The bottom panel of figure 2 compares the two models' predicted shares migrating *between* various combinations of sample tracts. This comparison shows that the restricted model's predictions miss significant variation in tract-to-tract flows, substantially over-predicting flows between some tract pairs and under-predicting flows between other pairs. These patterns suggest that allowing for rich heterogeneity in preferences over prices and locations *within* broad demographic groups is crucial if one's aim is to recover realistic patterns of substitutability between neighborhoods.

Figures 3 through 6 illustrate the flexibility of our specification across types and across neighborhoods graphically. Figure 3 shows a map of Los Angeles county for reference, and figures 4, 5 and 6 show spatial estimates of δ for three different types. Different types place very different relative values on the same location, which would be consistent with types making very different location decisions. This dramatic variation across people in the relative value of neighborhoods argues for an estimation approach that allows for many types, which is only possible with very a large data set such as ours.

3 Neighborhood Effects

In this section, we use confidential panel data from the Los Angeles Family and Neighborhoods Survey (LA FANS) study how neighborhoods impact child cognitive abilities. We do not take a stand on whether tract-by-tract variation we uncover in these neighborhood effects arise from differences in school quality, peer effects, or something else. The LA FANS study was designed specifically to investigate neighborhood influences on a variety of outcomes for families, adults, and children; see Pebley and Sastry (2011). The survey stratified 65 Census tracts using 1990 boundaries in Los Angeles County. Roughly 50 households in each Census tract were selected at random for inclusion in the survey. A randomly selected adult in the household was interviewed, as well as a randomly selected child. If the household had more than one child, a randomly selected sibling was also interviewed. Further, if the selected child's mother was in the household, she was interviewed as the primary caregiver. If she was absent, the actual primary caregiver was interviewed.

The LA FANS data has the advantage of sampling by Census tract, so that we observe many households within a small geographic region.[7] The LA FANS oversamples poor neighborhoods, but the 65 Census tracts are distributed across much of Los Angeles. 3,085 households were interviewed between 2000 and 2002 (wave 1), of which 1,242 were re-interviewed between 2006 and 2008 (wave 2). New households were admitted into the LA FANS sample in the second wave. Detailed information on the housing status (rentership versus ownership), family characteristics, and child outcomes were collected from respondents and Census

[7]This is in contrast with other geo-coded panel datasets such as the Panel Survey of Income Dynamics or the National Longitudinal Study of Youth.

tract information was collected in both waves.

We study two different cognitive skill measures as dependent variables. These measures are the child's score on Woodcock Johnson tests as described in Schrank, McGrew, and Woodcock (2001) for applied problems ("math") and passage comprehension ("reading"), tests used in many MTO studies. We restrict our sample to children who had valid measurements for both waves and we eliminate from our sample children with missing observations in some of our control variables.[8] This reduces our sample to $1,260$ for our math skill measure and $1,274$ for our reading skill measure.[9]

We compute measures of neighborhood value added in a manner that is analogous to a standard technique in the education literature for computing teacher value added. Following, for instance, Kane and Staiger (2008) and Chetty, Friedman, and Rockoff (2014) we work with the statistical model for the production of the change between periods $t-T$ and t in several child ability measures ($\Delta_{t-T}A_{i,j,t}$),

$$\Delta_{t-T}A_{i,j,t} = Z'_{i,j,t-T}\psi + v_{i,j,t} \quad ; \qquad v_{i,j,t} = T\mu_j + \epsilon_{i,j,t} \; , \tag{17}$$

where i indexes children, j indexes neighborhoods, t indexes time, $Z_{i,j,t-T}$ is a vector of observable child and family characteristics measured at time $t-T$, T is the time between LA FANS waves, μ_j is a causal (annualized) neighborhood "value-added" effect, and $\epsilon_{i,j,t}$ is an idiosyncratic child/family effect.

Notice that in the absence of any control variables, μ_j would govern the average change in child ability over time for children living in neighborhood j. Consistent with the value-added approach, splines of lagged values of the Woodcock Johnson test of letter-word identification and a behavioral problems index as described in (Peterson and Zill, 1986) are included as controls. Our other controls include variables covering family structure (number of children), language spoken, race, gender of child, parental cognitive ability (also captured by Woodcock Johnson tests), parental education, earnings and assets. We present descriptive statistics of our key dependent and independent variables in Table 2.

The key insight to the value-added approach is that parents' optimal neighborhood choice does not have to be uncorrelated with the observable control variables, including lagged child test scores, to produce unbiased estimates of neighborhood effects on child ability. Due to the presence of neighborhood fixed effects in equation (17), ψ is identified purely by

[8]Children that change locations between waves are assigned to the Census tract of their location in the first wave.

[9]A major reason for a lack of skill measurement in both waves is the child's age. Only children under 18 were administered the Woodcock Johnson tests. This means that only children who were under 18 in wave 2, i.e. aged 4 to 14 in wave 1 depending on the interview timing, would be included. Furthermore, new entrants to the survey would be disqualified since we only see their skills once.

within-neighborhood variation of $Z_{i,j,t-T}$ and $\Delta_{t-T}A_{i,j,t}$. Parents can select neighborhoods based on $Z_{i,j,t-T}$ and that will not bias estimates of ψ. For an unbiased estimate of μ_j, the error term $\epsilon_{i,j,t}$ must be uncorrelated with $Z_{i,j,t-T}$.[10] For our results to be misleading, selection into neighborhoods based on $\epsilon_{i,j,t}$ must account for a significantly larger share of observed differences in change in average ability across neighborhoods than selection into neighborhoods based on parental education, income and assets (Altonji, Elder, and Taber, 2005).

Following the teacher value added literature, we compute empirical Bayes estimates of neighborhood value added estimates $\widehat{\mu}_j$. The slope coefficients ψ are estimated in a first stage by regressing ability scores $\Delta_{t-T}A_{i,j,t}$ on $Z_{i,j,t-T}$ and a set of neighborhood fixed effects. Neighborhood value added measures are then computed in a second stage using,

$$\widehat{\mu}_j = \frac{1}{T}\Big(\frac{1}{N_j}\sum_{i\in j}\widehat{v}_{i,j,t}\Big)\left(\frac{\widehat{\sigma}^2_\mu}{\widehat{\sigma}^2_\mu + \widehat{\sigma}^2_\epsilon/N_j}\right) \tag{18}$$

In the above equation N_j is the number of observations in neighborhood j, $\widehat{\sigma}^2_\mu$ is the standard deviation of the estimated neighborhood fixed effects and $\widehat{\sigma}^2_\epsilon$ is the standard deviation of child outcomes after controlling for all Z terms and neighborhood effects. The first term in parentheses in the equation above is the average of the estimated residuals $\widehat{v}_{i,j,t} = A_{i,j,t} - Z'_{i,j,t-T}\widehat{\psi}$ within neighborhood j. The second term in parentheses shrinks this average toward zero as in Chetty, Friedman, and Rockoff (2014). This correction accounts for extra variation in estimated neighborhood effects arising from sampling uncertainty, i.e. small sample sizes in each neighborhood.[11]

Table 3 summarizes our regression results, showing model fit across a number of specifications. The outcome variable is the change in the relevant standardized test score between LA FANS waves. Overall, the neighborhood fixed effects and our full set of controls explain 50 percent of the variation in the change in math outcomes and 42 percent of the change in reading outcomes. As the first row of table 3 shows, neighborhood fixed effects alone explain 18-19 percent of the variation in test scores across children. Once we add splines for lagged child scores, specification 2, the regressions explain about 41-48 percent of the variation. We interact demographic information about the child with splines of the lagged test scores and with each other, specification 3, which boosts the $R2$ to 50% for reading and 57% for math. Information about the parents ability and demographics (specification 4) and

[10] This is equivalent to saying that parents can select neighborhoods based on the level of their child's ability and/or other variables in $Z_{i,j,t-T}$, but not on the portion of expected growth of child ability that is not forecasted by $Z_{i,j,t-T}$.

[11] The intuition for the correction is from the measurement error literature.

household income and assets if positive (specification 5, the full model) explain very little of the change in child outcomes, conditional on the other variables in the regression.[12]

Figure 7 shows how our estimated distribution of neighborhood value added changes with each of the specifications. The black line corresponds to specification 1, the regression with only neighborhood fixed effects. The dotted red line is for specification 2, the same as 1 but with splines in lagged child scores; the dashed-red line adds to that child characteristics (specification 3); the solid red line adds parental ability and demographics (specification 4); and the dashed blue line is the full model, specification 5, including income and assets. Consistent with the results in table 3, the neighborhood distributions do not change much once the model includes lagged test scores and child controls. Using our full specification, we estimate that the standard deviation of neighborhood value added accumulating between LA FANS waves accounts for about 5% of the cross-sectional variance in child ability.

In order to better understand our value-added measures, we correlate them with various neighborhood characteristics including percentages of black and hispanic, poverty rates and other economic characteristics, and school quality. The correlations are shown in Table 4. The size of the correlations are generally small but have the sign we expect, such as the positive correlation of income and negative correlation of unemployment. Our value added measures is also positively correlated with the quality of the attached public schools (specifically, measures of the schools' value added published by the L.A. Times[13]), though the fact that this correlation is also low suggests that most of our measured neighborhood effects are driven by mechanisms other than the quality of local schools.

The LA FANS data cover 65 of Los Angeles County's roughly 2000 Census tracts. To continue our analysis, we impute neighborhood value added estimates for the non-LA FANS tracts in Los Angeles by taking spatial moving averages of the LA FANS-based estimates. Specifically, we compute:

$$\widehat{\mu}_j = \frac{\sum\limits_{j' \in LAFANS} \phi\Big(\dfrac{dist(j,j')}{h}\Big)\ \widehat{\mu}_{j'}}{\sum\limits_{j' \in LAFANS} \phi\Big(\dfrac{dist(j,j')}{h}\Big)} \tag{19}$$

where $\phi()$ is a normal kernel, $dist(j, j')$ is the distance between the centroids of tracts j and j', and h is the bandwidth. To select a bandwidth, we repeatedly implement a leave-one-out jacknife version of this procedure within the LA FANS sample over a range of bandwidths

[12] All demographic variables and income and assets are measured during wave 1 of the LA FANS Survey.

[13] See *http://projects.latimes.com/value-added/* for details on how the school value-added measure is computed. We assign the elementary school that is closest in distance to the centroid of the Census Tract.

and select the bandwidth that minimizes the mean squared deviation of these spatial moving averages from tracts' actual value added estimates. We then apply the procedure to all tracts using this optimal bandwidth. The optimal bandwidth is just above one mile, illustrated by 8 for the passage comprehension value-added estimates.

Figure 9 shows the proximity of non-LA FANS Census tracts to the nearest LA FANS Census tract separately by poverty category. The solid red line displays the cumulative density function for tracts with a poverty rate greater than 30%. About 70% of these tracts in Los Angeles county are located within two miles of a tract sampled in the LA FANS data, and the modal tract in this poverty category is located less than one mile from an LA FANS tract. Reflective of LA FANS' oversampling of poor tracts, on average low-poverty Census tracts are farther from an LA FANS tract.

4 Reconciling Large Neighborhood Effects with MTO

Our finding of large "neighborhood effects" is squarely in line with an earlier literature that estimates these effects: See Leventhal and Brooks-Gunn (2000) and Durlauf (2004) for recent surveys. While these studies typically attempt to account for selection issues,[14] the fact that individuals endogenously sort into neighborhoods leaves open the possibility of non-causal explanations for these patterns.[15]

Recognizing the limitations of observational studies, the literature on neighborhood effects has devoted considerable attention recently to the "Moving to Opportunity" randomized experimental intervention. Moving to Opportunity was a randomized control trial begining in the 1990s that randomly assigned a group of households eligible to live in low income housing projects in five U.S. cities to three different groups; (i) a treatment group that received a Section 8 housing voucher that in the first year could be applied only in Census tracts with a poverty rate under 10% and could be applied unconditionally thereafter, (ii) a second treatment group that received a comparable Section 8 housing voucher with no location requirement attached, and (iii) a control group that received no voucher. Summarizing the medium to long term impacts of MTO, Sanbonmatsu, Kling, Duncan, and Brooks-Gunn (2006), Kling, Liebman, and Katz (2007) and others show that on average the MTO treatment successfully reduced exposure to crime and poverty and improved the mental health

[14] For example, Cutler and Glaeser (1997) study the impact of segregation on outcomes of African-Americans using topographical features of cities as instruments for location choice and Aaronson (1998) measures neighborhood effects by studying outcomes of siblings at least three years apart in age after a move.

[15] See Aaronson (1998) for examples of instruments used by other researchers in this field and their potential limitations.

of female children, but failed to improve child ability, educational attainment, or physical health.[16]

But do the MTO results prove that neighborhood effects are small? Perhaps not. Suppose there is variation in neighborhood value-added in tracts with a poverty rate under 10%; and, suppose that rents are higher for tracts with greater value-added. Once households receive a voucher to live in a tract with pverty rate under 10%, they must decide whether to move to a high-rent, high-value-added tract or a low-rent, low-value-added tract. Figure 10 gives a stylized graphical illustration of the range of possible outcomes after an MTO-style intervention. In both panels, the x-axis represents neighborhood value-added; the y-axis represents housing rent; the solid black line shows the set of available combinations for the high-poverty neighborhoods; the dashed line shows the set of available combinations for the low-poverty neighborhoods; and the red lines show indifference curves.[17] The top panel shows one possible outcome from MTO: As households move from high-poverty to low-poverty tracts via the MTO rent subsidy, their rent falls and their child value-add rises. The bottom panel shows a case where child value-added falls after the MTO rent subsidy. Ultimately, the change in child outcomes after the rent subsidy is received depends on ideas from classic microeconomics: Changes to the slope of the budget line, and income and substitution effects.

Further, casual inspection of our data suggests relative prices, income and substitution effects may be of first order importance. Table 5 reports estimates from descriptive hedonic regressions our estimates of neighborhood value-added measures to median monthly housing rents from the 2000 Decennial Census for each Census tract. The first column reports regressions for only the 62 LA FANS tracts; the second column reports the same results for all 1,748 Census tracts in Los Angeles county in our neighborhood-choice study after we have imputed $\widehat{\mu}_j$ using equation (19) for all tracts; and the third column is the same as the second but it also includes basic demographic information in the regression.

These regressions are suggestive that the neighborhood rent gradient with respect to child ability value added is steeper in low poverty Census tracts than in high poverty Census tracts. Consider two otherwise identical Census tracts, but with one tract offering a one standard deviation increase to change in Math ability than the other over the course of 10 years.[18] According to estimates from the third column of table 5, this would be associated with a negligible decrease in monthly rent of $4 for tracts with a poverty concentration between 10%

[16] Recent work by Chetty, Hendren, and Katz (2015) argues that MTO positively affected adult wages.

[17] Households dislike housing rent and like value added, so households are best off in the south-east corner of the graph.

[18] Referring to the top panel of Figure 9, think of the first tract as +0.05 and the second tract as -0.05 for neighborhood value added per year.

and 25%. For tracts with a poverty concentration of less than 10%, the implied difference in monthly rent is a staggering $659. If a household wants to move into a low poverty neighborhood with high neighborhood value-added, estimates in table 5 suggest this will require much higher monthly rent than moving into a low poverty neighborhood with low child value-added.

Figure 11 visually tells a similar story. The figure plots neighborhood value added against median monthly rent for three groups of Census tracts: Low poverty concentration (0-10%), middle (10-25%), and high poverty (25% and above). These figures show how the relative price of neighborhood quality changes with tract poverty rates. The change in rent associated with an increase in neighborhood quality is greatest in low poverty areas; that is, the slope of the green line (low poverty) is greater than the slope of the blue line (middle), which is greater than that of the red line (high poverty). At best, in high poverty areas, child value added may be unpriced.

Even though neighborhoods with high value-added are relatively expensive in low poverty tracts, households may be willing to pay to live in those neighborhoods conditional on receiving a large enough rent subsidy. Therefore, within the context of our full model, to understand the impact of a rent subsidy program such as MTO on neighborhood choice (and thus child outcomes), we need to understand how utility of each neighborhood varies with rent. Denote as $\tilde{\delta}_{j\tau}$ our estimate of indirect utility of neighborhood j for given type τ. We specify that $\tilde{\delta}_{j\tau}$ is a linear function of rent, observables characteristics of tract j, $\mathcal{O}_j$, and unobserved characteristics of tract j, ζ_j

$$\tilde{\delta}_{j\tau} \;=\; -\alpha_\tau \cdot rent_j \;+\; \lambda_\tau \cdot \mathcal{O}_j \;+\; \zeta_j \tag{20}$$

α – the rate at which indirect utility varies with rents – in equation (20) cannot be estimated using OLS because equilibrium rents will almost certainly be correlated with unobserved (but valued) characteristics of neighborhoods, ζ_j. We use an IV approach to estimate (20) that is common in the IO and Urban literature, for example Bayer, Ferreira, and McMillan (2007). In $\mathcal{O}$ we include characteristics of the housing stock 0-5 miles from tract j and our instruments are characteristics of the housing stock 5-20 miles from the tract. These instruments are assumed to affect equilibrium rent in j but do not directly affect δ_j.

We find remarkable variation in our estimates of α by type. We summarize this variation by reporting the average value of α by initial Census tract of residence for the people in our FRBNY Consumer Credit Panel / Equifax estimation sample. This average value of α varies by Census tract because the mix of types varies by tract. We restrict our attention to tracts with a poverty rate less than 40%. There are tracts in our sample with higher poverty rates, but as Figure 12 shows, the number of types represented in each tract falls dramatically in neighborhoods above a 40% poverty rate. Figure 13 shows our estimates of the variation in the average value of α by poverty rate of initial tract of residence. The figure shows that people living in high-poverty tract areas are, on average, about twice as sensitive to changes in rent as people living in the lowest poverty tract areas.

Table 5 and figures 11 and 13 foreshadow our results. Figure 13 suggests the types of people currently living in high poverty tract areas are quite sensitive to the level of rent; and, Figure 11 and Table 5 suggests the relative price of a high-value-added neighborhood in a low-poverty-rate neighborhood is much greater than that of a high-poverty-rate neighborhood. It seems quite possible that child outcomes may not change – or might even worsen – if we subsidize people to move from high poverty neighborhoods to low poverty neighborhoods without further restricting which low poverty neighborhood they move to. If this were indeed the case, it would reconcile the apparent contradiction of large neighborhood effects in the observational literature and small experimental results of MTO.

Given our type-specific estimates of α, we conduct simulation experiments to better understand the implications of the MTO experiment. Specifically, we use our estimated model to simulate location sequences under several policy scenarios, restricting analysis to the households in our sample likely to have been eligible for MTO had they lived in an MTO area at the time of the experiment. Our three scenarios are as follows:[19]

- (Baseline) No subsidies or vouchers.
- (MTO-A) MTO style vouchers. Households who move to a Census tract with a poverty rate under 10% at $t = 1$ receive a Section 8 housing voucher that may be used in perpetuity. For any neighborhood j that qualifies, we set the utility of that neighborhood for type τ after the voucher is received equal to the original estimate of the utility of that neighborhood, $\tilde{\delta}_{j\tau}$, plus α times the voucher amount. The annual voucher we use is \$6,000, which we set such that the average MTO-eligible household can rent a 2-bedroom unit costing \$766 per month after spending 30% of monthly income.[20]
- (MTO-B) Randomly assigned poverty reduction. Assign households to neighborhoods randomly according to the distribution of neighborhood poverty-rates that arises under

[19]Our simulations target households residing at $t = 0$ in a Census tract with at least 500 public housing units. Alternative targeting rules (results not shown) targeting eligibility to residents of tracts with very high poverty rates and/or rates of public assistance yield similar results.

[20]Our calculation is $\$6,000 \approx 12\,[\$766 - 0.30\,(\$10,000/12)]$, where \$10,000 is mean household income of the MTO-eligible population as computed by Galiani, Murphy, and Pantano (2015) and \$766 is the "payment standard" (max voucher amount) for a 2-bedroom apartments in Los Angeles in 2000.

scenario MTO-A.[21]

Figure 14 shows the simulated distribution of the poverty rate of locations chosen in all periods of the simulation in our baseline case and in the MTO-A simulations. Simulations last for 18 years following implementation of the MTO policy. The distributions that are shown are for only the households that took up a voucher in the MTO-A simulations.[22] Comparing the black solid line (baseline) with the blue dashed line (MTO-A), our simulations find that most of the people induced by MTO to move to a low poverty neighborhood choose to remain in a low poverty neighborhood for an extended period of time. Among all households that took up a voucher in the MTO-A simulation, only 12% of person-years are spent in a neighborhood with a poverty rate greater than 10%; whereas in the baseline, 74% of person-years of those households are spent residing in those neighborhoods.

To summarize the expected impact on child ability of this reduction to poverty exposure, we compute an expected measure of accumulated neighborhood value-added exposure for a given individual i under government policy p as,

$$\widehat{\mu}_{i,p}^{TOT} = \frac{1}{S}\sum_{s=1}^{S}\sum_{t=1}^{\mathcal{T}}\widehat{\mu}_{\ell(i,t,s,p)} \tag{21}$$

where $\ell(i,t,s,p)$ is the location chosen by individual i in year t under policy p and for given simulation draw s and $\mathcal{T} = 18$ is the number of years in the simulation. For each type, we run 10,000 simulations, yielding a total of 1.44 million simulations for each policy experiment. If, as suggested by Chetty and Hendren (2015), neighborhood effects are additive over time in the child ability production function (i.e. there are no complementarities across time periods) and neighborhood quality affects children equally at all ages, then these measures will characterize actual total neighborhood contributions to child ability. If child investments exhibit dynamic complementarities and early childhood investments are especially productive as in Cunha, Heckman, and Schennach (2010), these measures will understate neighborhoods' long-term contributions to child ability. In either case, we view these measures as useful summaries for characterizing policies' impacts.

Figure 15 plots cumulative policy impacts on reading-scores value-added (top panel)

[21] Specifically, the procedure is; (1) pool the set of MTO-A simulated Census tract choices and the unconditional list of sample Census tracts. (2) Estimate a probit model predicting the probability that a record comes from the simulated data using only tract-poverty-rate categories as explanatory variables, and obtain the predicted probability p_j (propensity score) that a record from tract j comes from the simulated data. (3) Draw MTO-B simulated locations from the full set of Census tract with probability $Pr(j) = \frac{1}{J}\Big(\frac{p_j}{1-p_j}\Big)\Big(\frac{1-\overline{p}}{\overline{p}}\Big)$.

[22] In our MTO-A simulations, 70% of the population eligible to receive a voucher accept it; the actual MTO take-up rate in Los Angeles was 67%.

and math-scores value-added (bottom panel), relative to the baseline scenario. The policy impacts are plotted by year after the implementation of the MTO experiment. In both panels, the dot-dashed blue line shows the predicted impact of MTO-B and the solid black line shows the predicted impact of MTO-A. As with figure 14, the data are from the sample of households that take up a voucher in the MTO-A simulation. One test of the accuracy of our research program is to see if MTO-A's predicted impact relative to baseline is consistent with the zero-impact of the MTO experiment on children's math and reading ability. Indeed, the solid line on both plots shows an approximately zero predicted impact for the MTO-A scenario.

MTO's zero impact has been cited as evidence that the "average treatment effect" (ATE) of lower poverty neighborhoods on children's cognitive ability is negligible. An advantage of our framework is that we can directly compute this ATE by studying the impact of the MTO-B scenario that randomly assigns households to lower poverty neighborhoods. The MTO-B simulations find that, accumulated over a full 18-year childhood, the poverty reduction generated by MTO would improve both math and reading scores by about 0.2 standard deviations if low-poverty neighborhoods were assigned at random. These are substantial impacts, equivalent to closing about 20% of the black/white achievement gap according to Yeung and Pfeiffer (2009). Taken together, the MTO-A and MTO-B results suggest that MTO-subsidized households selected into especially low value-added tracts among the set of eligible low-poverty tracts.

5 Conclusion

In this paper, we use two new rich data sets to understand how households choose neighborhoods and the impact of neighborhoods on child ability. We find considerable heterogeneity in the population in the utility of different neighborhoods; and we also show meaningful variation in the impact of neighborhoods on child ability as measured by test scores. We also show that the utility of households residing in high-poverty neighborhoods, on-average, is much more sensitive to rental prices than the utility of households residing in low-poverty neighborhoods. This last result reconciles two results in the literature on child outcomes that seem contradictory: The existence of large effects of neighborhoods on changes to child ability, and the overall lack of improvement of children in the MTO experiment. Counterfactual simulations of our model of neighborhood choice strongly suggest that parents receiving vouchers in the MTO experiment moved to the lowest-cost, lowest-value-added neighborhoods among the eligible set. If parents had randomly chosen low-poverty neighborhoods after receiving a voucher, our analysis suggests their children would have shown a remarkable improvement in ability.

References

Aaronson, D. (1998): "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes," *The Journal of Human Resources*, 33(4), 915–946. 2, 14

Aliprantis, D. (2015): "Assessing the Evidence on Neighborhood Effects from Moving to Opportunity," Federal Reserve Bank of Cleveland Working Paper. 3

Altonji, J. G., T. E. Elder, and C. R. Taber (2005): "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 113(1), 151–184. 12

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. 3, 16

Bayer, P., R. McMillan, A. Murphy, and C. Timmins (2015): "A Dynamic Model of Demand for Houses and Neighborhoods," Duke University Working Paper. 1, 4, 5, 8

Bhutta, N., and B. J. Keys (2015): "Interest Rates and Equity Extraction During the Housing Boom," Kreisman Working Papers Series in Housing Law and Policy No. 3. 8

Bishop, K. C. (2012): "A Dynamic Model of Location Choice and Hedonic Valuation," Working Paper, Washington University in St. Louis. 6

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. 5

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," . 7

Brown, M., S. Stein, and B. Zafar (2013): "The Impact of Housing Markets on Consumer Debt: Credit Report Evidence from 1999 to 2012," Federal Reserve Bank of New York Staff Report No. 617. 8

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014): "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates," *American Economic Review*, 104(9), 2593–2632. 2, 11, 12

Chetty, R., and N. Hendren (2015): "The Impacts of Neighborhoods on Intergenerational Mobility: Childhood Exposure Effects and County-Level Estimates," Harvard University Working Paper. 18

Chetty, R., N. Hendren, and L. F. Katz (2015): "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," National Bureau of Economic Research Working Paper 21156. 3, 15

Clampet-Lundquist, S., and D. S. Massey (2008): "Neighborhood Effects on Economic Self-Sufficiency: A Reconsideration of the Moving to Opportunity Experiment," *American Journal of Sociology*, 114(1), 107–143. 3

Cunha, F., J. J. Heckman, and S. M. Schennach (2010): "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 78(3), 883–931. 18

Cutler, D. M., and E. L. Glaeser (1997): "Are Ghettos Good or Bad?," *The Quarterly Journal of Economics*, 112(3), 827–872. 1, 14

Davis, M. A., and E. M. Foster (2005): "A Stochastic Dynamic Model of the Mental Health of Children," *International Economic Review*, 46(3), 837–866. 3

Durlauf, S. N. (2004): *Handbook of Regional and Urban Economics*chap. Neighborhood Effects, pp. 2174–2230. Elsevier B.V. 1, 14

Galiani, S., A. Murphy, and J. Pantano (2015): "Estimating Neighborhood Choice Models: Lessons from a Housing Assistance Experiment," University of Maryland Working Paper. 3, 17

Gallagher, J., and D. A. Hartley (2014): "Household Finance after a Natural Disaster: The Case of Hurricane Katrina," Federal Reserve Bank of Cleveland Working Paper No. 14-06. 8

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. 6

Kane, T. J., and D. O. Staiger (2008): "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation," National Bureau of Economic Research Working Paper 14607. 2, 11

Kennan, J., and J. R. Walker (2011): "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 79(1), 211–251. 1, 4, 5

Kling, J. R., J. B. Liebman, and L. F. Katz (2007): "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75(1), 83–119. 3, 14

Leventhal, T., and J. Brooks-Gunn (2000): "The Neighborhoods They Live in: The Effects of Neighborhood Residence on Child and Adolescent Outcomes," *Psychological Bulletin*, 126(2), 309–337. 1, 14

Ludwig, J., G. J. Duncan, L. A. Gennetian, L. F. Katz, R. C. Kessler, J. R. Kling, and L. Sanbonmatsu (2013): "Long-Term Neighborhood Effects on Low-Income Families: Evidence from Moving to Opportunity," *American Economic Review*, 103(3), 226–31. 3

Pebley, A. R., and N. Sastry (2011): "Los Angeles Family and Neighborhood Study (L.A. FANS)," RAND Corporation, Restricted File, Version 2.5. 10

Peterson, J. L., and N. Zill (1986): "Marital Disruption, Parent-Child Relationships, and Behavior Problems in Children," *Journal of Marriage and Family*, 48(2), 295–307. 11

Pinto, R. (2014): "Selection Bias in a Controlled Experiment: The Case of Moving to Opportunity," University of Chicago Working Paper. 3

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55(5), 999–1033. 6

Sanbonmatsu, L., J. R. Kling, G. J. Duncan, and J. Brooks-Gunn (2006): "Neighborhoods and Academic Achievement: Results from the Moving to Opportunity Experiment," *Journal of Human Resources*, 41(4), 649–691. 3, 14

Schrank, F. A., K. S. McGrew, and R. W. Woodcock (2001): *Woodcock-Johnson III Assessment Service Bulletin Number 2*Riverside Publishing, Itasca, IL. 11

Todd, P. E., and K. I. Wolpin (2006): "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility," *American Economic Review*, 96(5), 1384–1417. 3

Yeung, W.-J. J., and K. M. Pfeiffer (2009): "The black-white test score gap and early home environment," *Social Science Research*, 38(2), 412–437. 19

Table 1: Annual Cross-Tract Migration Rates

	% Moving Annually
All	8.4
Race:	
White	8.0
Black	8.9
Hispanic	8.6
Other	8.9
Age:	
Under 30	11.2
30-44	9.3
45-54	7.4
55-64	5.6
65+	3.2

Note: Race refers to racial plurality of the Census block of residence.

Table 2: Descriptive Statistics

	Mean	S.D.	Obs.
Dependent Variables (LAFANS wave 2)			
Math (z-score)	0.153	1.249	1260
Reading (z-score)	-0.195	1.199	1274
Control Variables (LAFANS wave 1)			
Math (z-score)	0.197	1.074	1357
Reading (z-score)	0.229	1.116	1357
Hispanic	0.581		1357
Black	0.120		1357
Male	0.520		1357
Parental IQ	87.851	15.416	1357
Parent dropout	0.326		1357
Parent high school	0.210		1357
Parent some college	0.292		1357
Parent bachelor	0.102		1357
Parent graduate	0.064		1357
Log earnings	9.731	3.052	1283
Log assets	2.727	1.911	1076

Table 3: Fit of Value Added Models

		Math		Reading	
Specification	Controls	R2	Adj. R2	R2	Adj. R2
(1)	Neighborhood Fixed Effects	0.177	0.136	0.186	0.146
(2)	+ Splines in Lagged Child Scores	0.481	0.446	0.412	0.373
(3)	+ Splines interacted w/ Child Controls	0.570	0.514	0.503	0.440
(4)	+ Parent Ability and Demographics	0.581	0.524	0.516	0.451
(5)	+ Lagged Income and Assets	0.583	0.525	0.519	0.453
(6)	Optimal FIC	0.502	0.465	0.423	0.378

Table 4: Descriptive Correlations

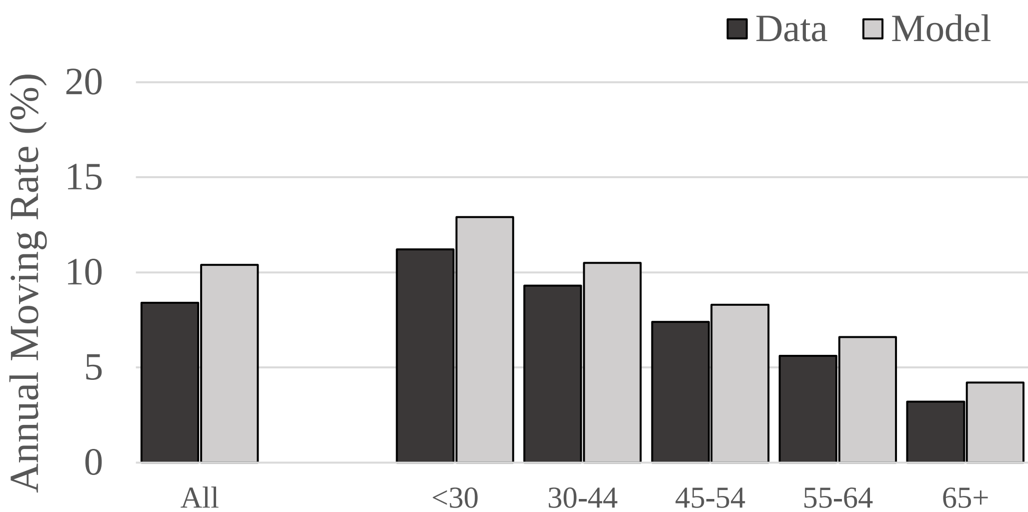
Neighborhood Characteristics	Math VA	Reading VA
Share hispanic	-0.216	-0.262
Share black	0.036	0.016
Average household income	0.090	0.147
Welfare	-0.116	-0.119
Poverty	-0.067	-0.077
Unemployment	-0.073	-0.050
School quality	0.100	0.104
Share of elderly	0.077	0.128

Table 5: Descriptive Hedonic Regression

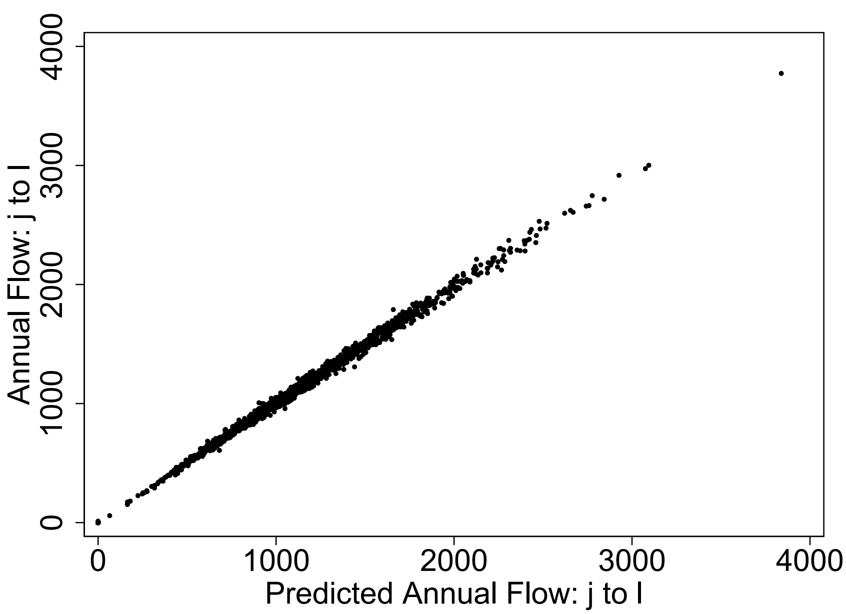
Controls	LAFANS	All	All
Value Added	99.31	-442.6	-616.1**
	(1,563)	(329.2)	(313.1)
Value Added x (Poverty 10%-25%)	413.7	1,187**	612.0
	(2,237)	(462.1)	(439.6)
Value Added x (Poverty < 10%)	-3,980	1,533***	1,275***
	(2,469)	(447.2)	(424.9)
Poverty 10%-25%	175.6**	153.8***	76.82***
	(71.47)	(13.40)	(13.71)
Poverty <10%	500.5***	462.8***	270.8***
	(77.75)	(13.86)	(18.41)
Pct. Hispanic			-348.3***
			(24.29)
Pct. Black			-289.4***
			(33.90)
Constant	580.2***	578.3***	860.4***
	(47.47)	(10.16)	(21.26)
Observations	59	1,916	1,916
R-squared	0.448	0.401	0.464

Note: The dependent variable in each regression is the Census tract median rent for the year 2000. The value added measures included in this table are for the Woodcock Johnson "applied problems" component. For column (1), the sample is restricted to Census tracts covered by the LAFANS with sufficiently many children sampled to compute neighborhood value-added estimates. Columns (2) and (3) include all of the Census tracts from Los Angeles county that we include in our neighborhood-choice analysis. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure 1: Model Fit

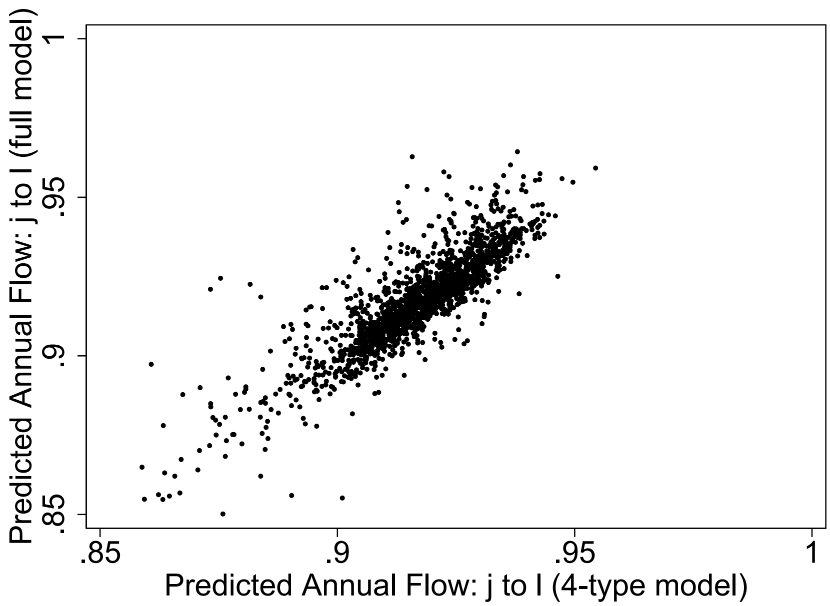


(a) Annual Moving Rates

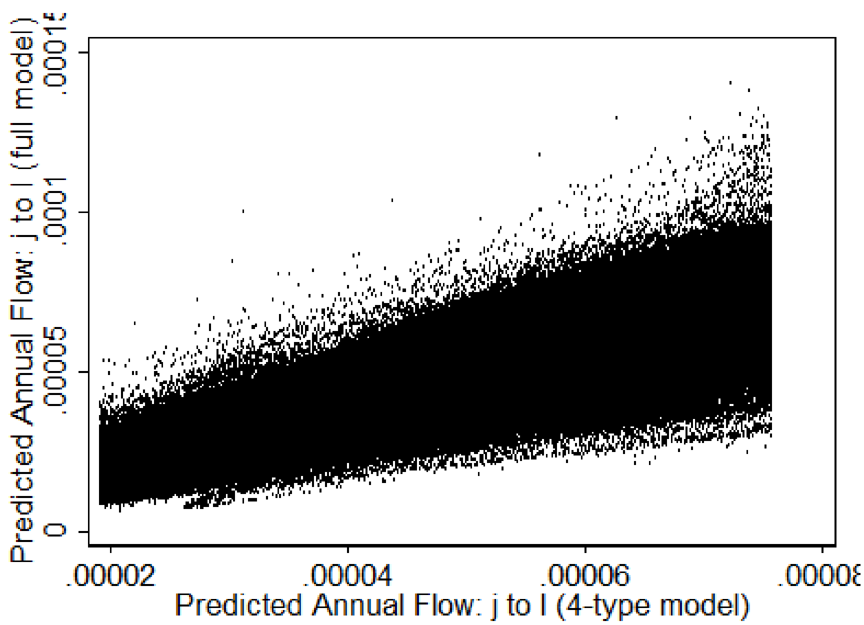


(b) Annual Tract-toTract Flows

Figure 2: Predictions of the Full Model vs. the Restricted (4-Type) Model



(a) Annual fraction of j's residents not moving (i.e. $\ell = j$)



(b) Annual fraction of j's residents moving to $\ell \neq j$

Figure 3: Los Angeles County

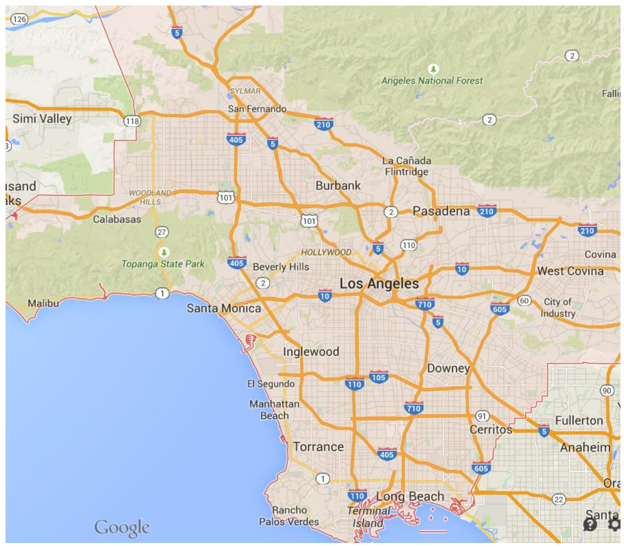


Figure 4: Example Spatial Spline - Type 93

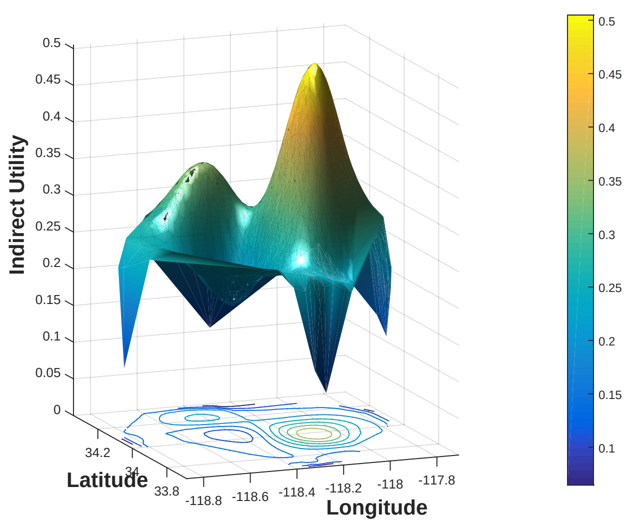


Figure 5: Example Spatial Spline - Type 119

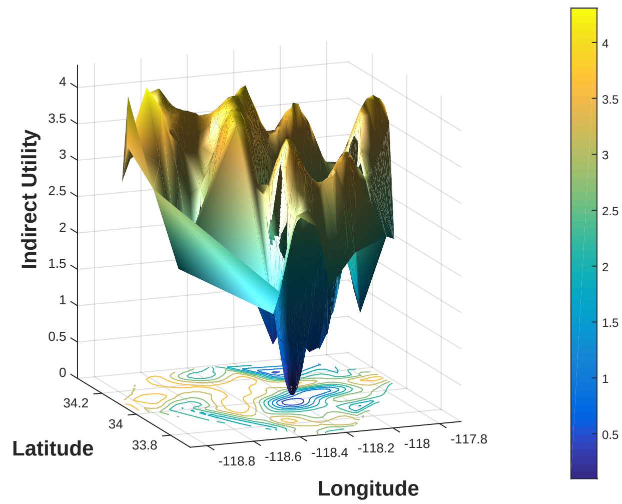


Figure 6: Example Spatial Spline - Type 129

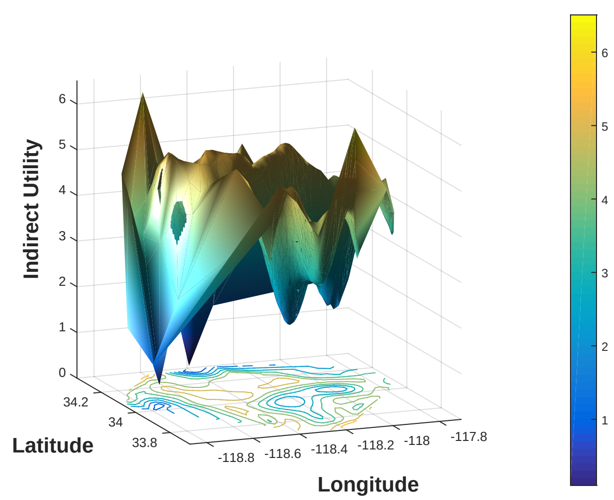
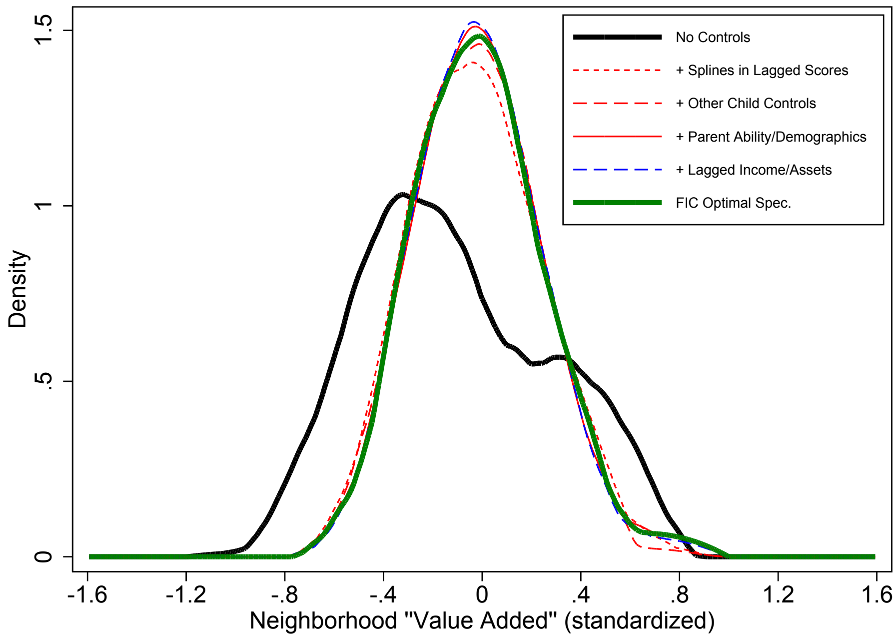
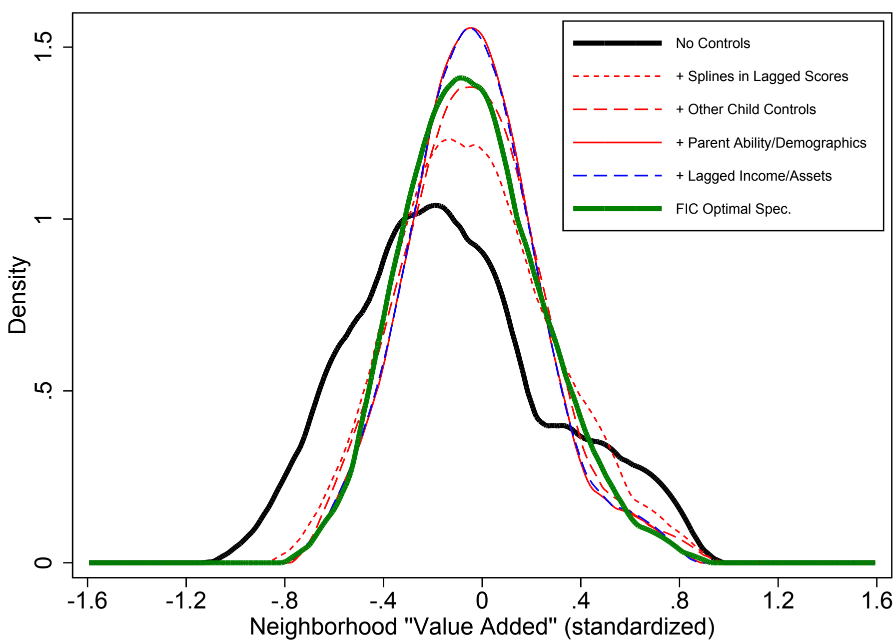


Figure 7: Estimated Neighborhood Value-Added Distributions by Specification



(a) Math



(b) Reading

Figure 8: Bandwidth Selection Criteria

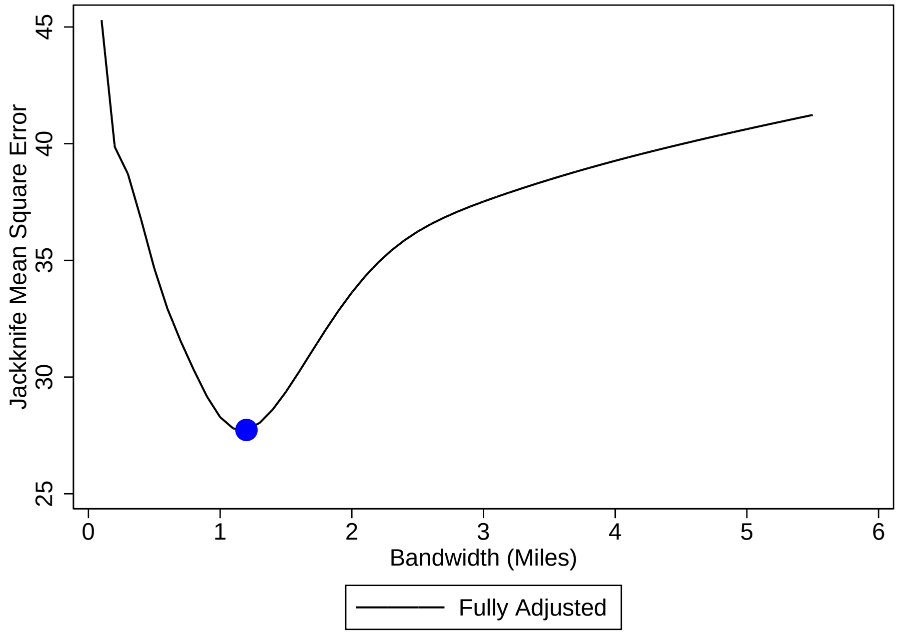


Figure 9: Los Angeles Census Tract Locations Relative to LA FANS Census Tracts

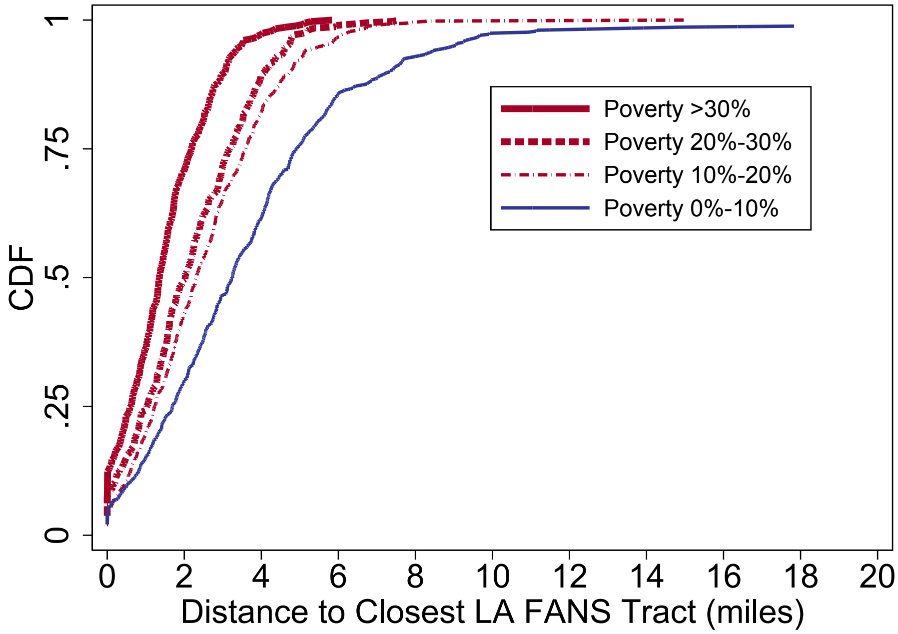
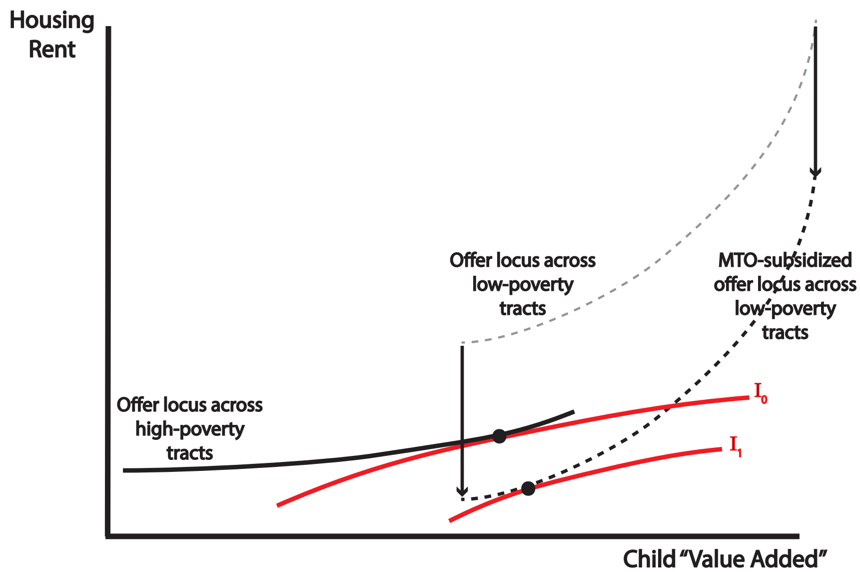
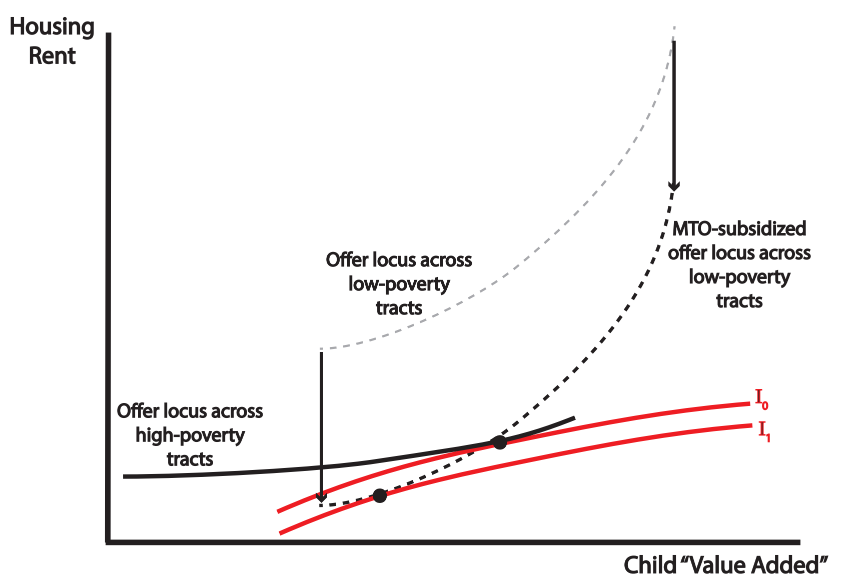


Figure 10: MTO's Predicted Effect on Child Value-Added when the Hedonic Rent/Value-added gradient is steeper in low-poverty areas than high poverty areas



(a) Scenario where MTO increases child outcomes



(b) Scenario where MTO decreases child outcomes

Figure 11: Rent, Neighborhood Value Added, and Poverty Rates

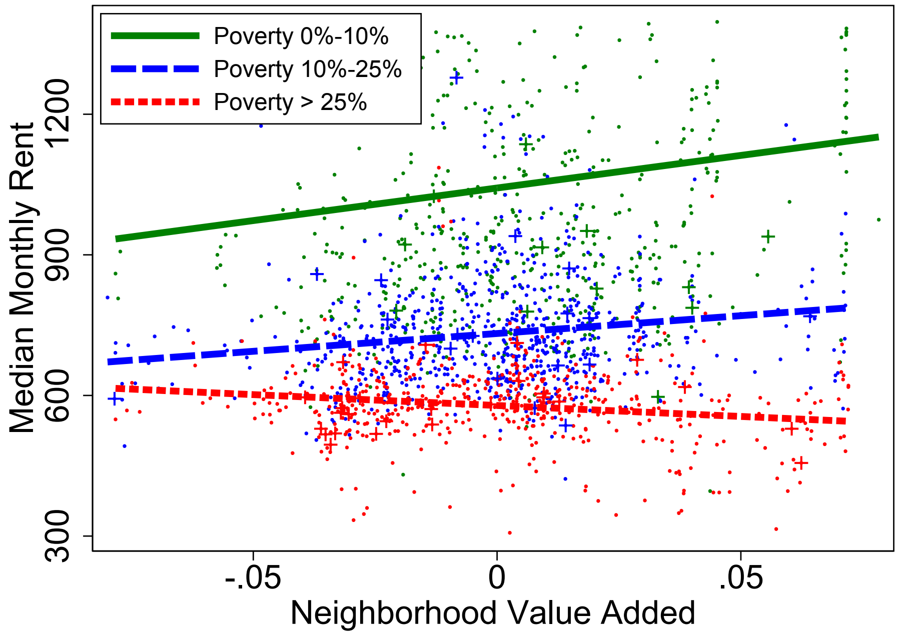


Figure 12: Counts of α

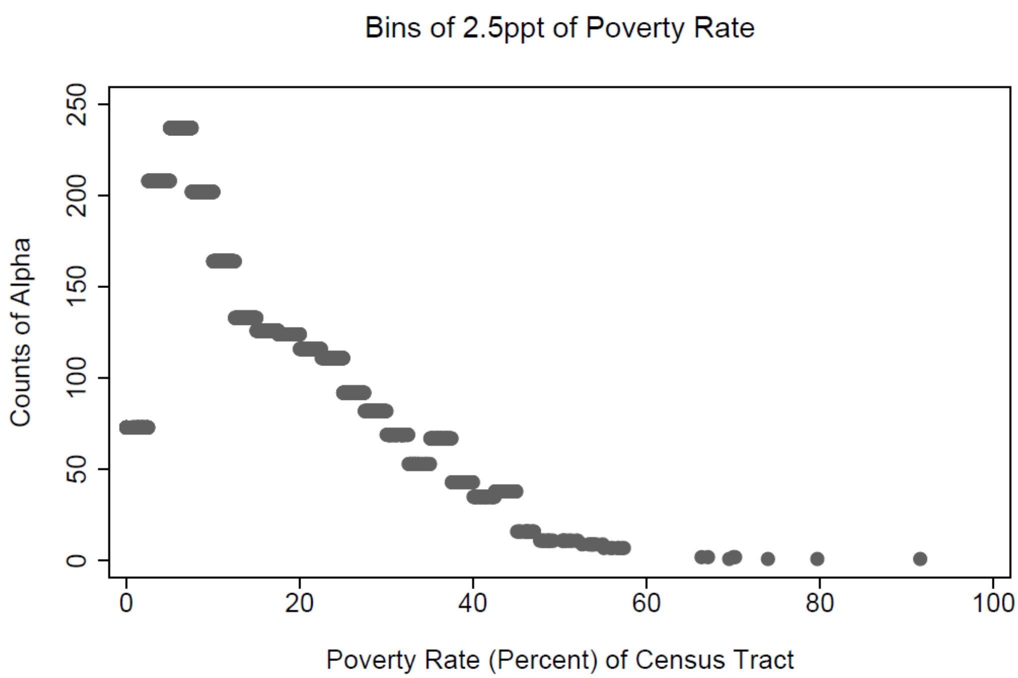


Figure 13: Average Value of α by Poverty Rate of Census Tract

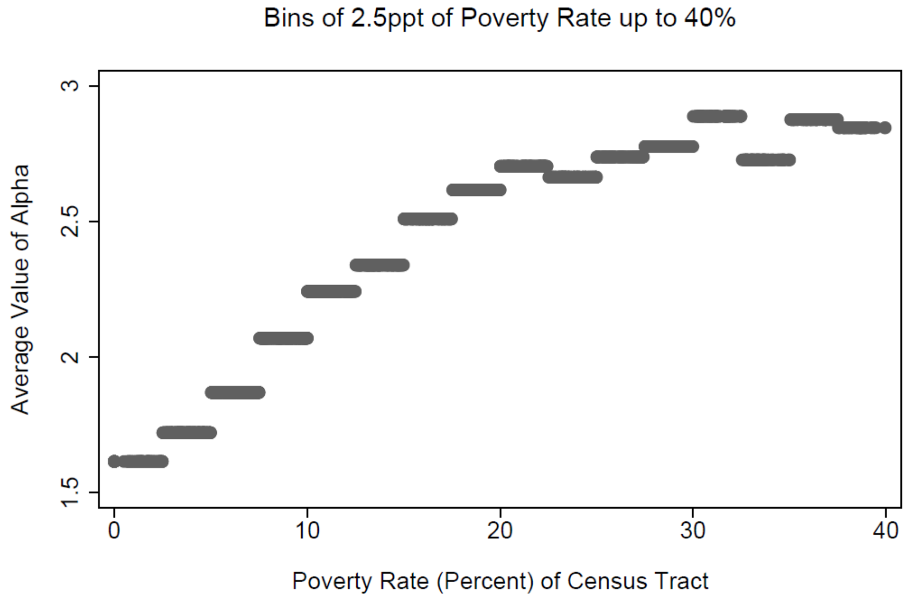


Figure 14: Distribution of Residents by Poverty Rate of Census Tract, Baseline and MTO

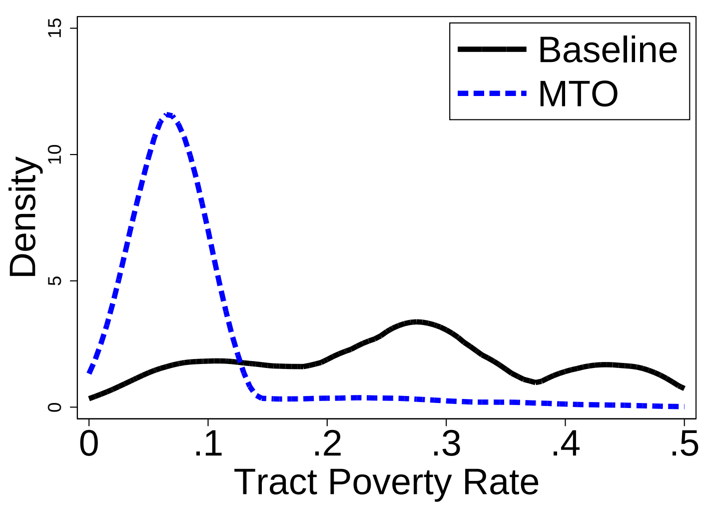
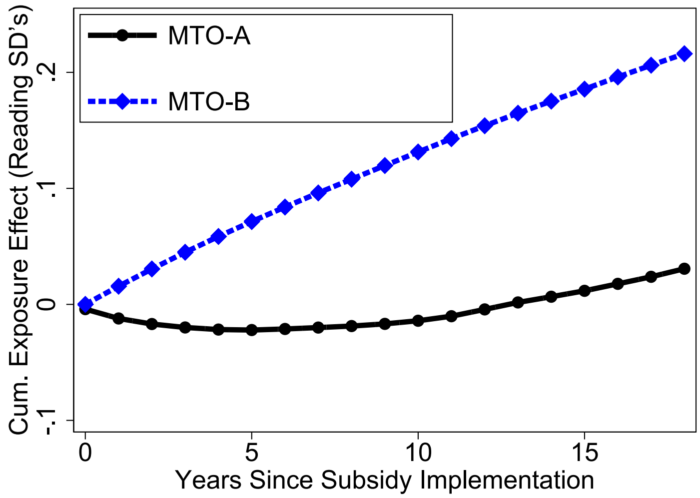
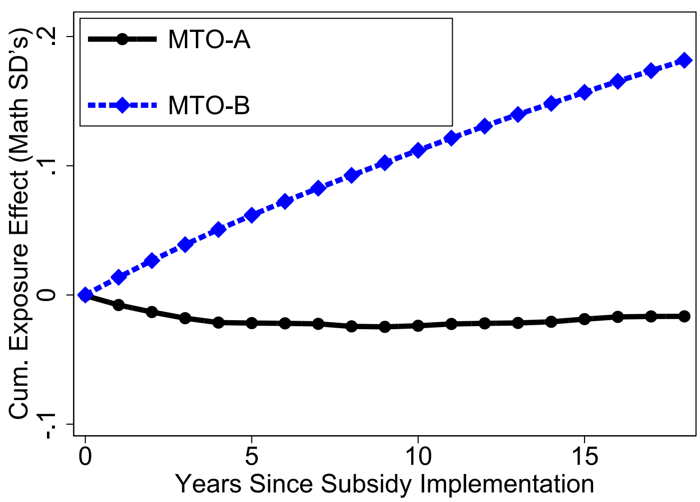


Figure 15: Counterfactual Simulations



(a) Reading Scores



(b) Math Scores