working paper

11 22

Assessing the Evidence on Neighborhood Effects from Moving to Opportunity

by Dionissi Aliprantis

FEDERAL RESERVE BANK OF CLEVELAND

Working Paper 11-22 September 2011

Assessing the Evidence on Neighborhood Effects from Moving to Opportunity

by Dionissi Aliprantis

This paper presents a new perspective on results from the Moving to Opportunity (MTO) housing mobility program. Building on recent developments in the program evaluation literature, this paper defines several treatment effect parameters and then estimates and interprets these parameters using data from MTO. The evaluation framework is used to make a clear distinction between the interpretation of Intent to Treat (ITT) and Treatment on the Treated (TOT) parameters as program effects and Local Average Treatment Effect (LATE) parameters as neighborhood effects. This distinction helps to clarify that results from MTO are only informative about a small subset of neighborhood effects of interest. Tests for instrument strength show that MTO induced large changes in neighborhood poverty rates. However, it is also shown that MTO induced remarkably little variation in many of the other neighborhood and school characteristics believed to influence outcomes and that much of this variation was confined to the tails of these characteristics' national distributions. Consistent with prevailing theories of neighborhood effects, labor market and health outcomes improved when households moved to neighborhoods with characteristics at or above the national median. The evidence suggests housing mobility programs designed to induce moves to neighborhoods with characteristics in addition to or in lieu of low poverty might induce larger effects than MTO, and results point to the investigation of heterogeneity in program effects from MTO as a fruitful direction for future research.

Keywords: Moving to Opportunity (MTO), Local Average Treatment Effect (LATE), Treatment-on-the-Treated (TOT), Housing Mobility Program, Neighborhood Effect, Concentrated Poverty, Segregation, Social Experiment.

JEL Classification Numbers: C30, H50, I38, J10, R00.

Dionissi Aliprantis is at the Federal Reserve Bank of Cleveland. He can be reached at (216)579-3021 or dionissi.aliprantis@clev.frb.org. The author thanks Francisca G.-C. Richter, Jeffrey Kling, Becka Maynard, Juan Pantano, and Ruby Mendenhall for helpful comments. Mary Zenker provided valuable research assistance, and Paul Joice at HUD has been extremely helpful. The research reported here was supported in part by the Institute of Education Sciences, U.S. Department of Education, through Grant R305C050041-05 to the University of Pennsylvania.

1 Introduction

"The problem of the Twentieth Century" has yet to be resolved. For nearly every outcome of importance, the distributions of blacks and whites in the United States are dramatically different. Persistent racial gaps can still be found whether one looks at educational outcomes such as test scores (Reardon (2008)) and attainment (Heckman and LaFontaine (2010)) or later outcomes such as earnings (Neal and Johnson (1996), Keane and Wolpin (2000)), incarceration (Pettit and Western (2004)), and exposure to violence (Figure 1, NCHS (2009)).

Although these outcomes have received much attention from social scientists, the mechanisms maintaining racial gaps are not well understood. One prominent theory proposes that effects from living in a poor, segregated, and socially isolated neighborhood can explain these differences in outcomes. The seminal work in Wilson (1987) presents empirical evidence on a trend of increasing concentration of urban poverty in the US, especially in predominantly African American neighborhoods. Indicative of this trend is that the number of people living in census tracts with poverty rates of 40% or more increased from 4.1 to 8.0 million between 1970 and 1990 (Ludwig et al. (2001)).[1] Wilson (1987) posits that growing up in a neighborhood of such concentrated poverty tends to have negative effects on outcomes.[2]

Policy makers have looked to housing mobility programs as a way to mitigate the adverse effects of concentrated poverty and segregation ever since the promising results of the Gautreaux program. The Gautreaux program relocated public housing residents in Chicago through housing vouchers in a quasi-random manner. Those who moved to low-poverty suburbs through Gautreaux had much better education and labor market outcomes than those who moved to city neighborhoods (Rosenbaum (1995)). Moving to Opportunity (MTO) was a housing mobility experiment conducted in five US cities seeking to replicate the quasi-experimental results from Gautreaux. Households living in high-poverty neighborhoods were allowed to enter a lottery for housing vouchers. In a tremendous disappointment to researchers and all those hoping to live in a society with equality of opportunity, MTO did not reproduce the beneficial effects found in Gautreaux.

This paper presents a new perspective on the interpretation of results from MTO, especially as they relate to neighborhood effects. In addition to studying the effects of specific housing voucher policies, researchers have interpreted estimates of the effects of moving through MTO as neighborhood effects. This paper argues that such an interpretation of results from MTO conflates program effects with neighborhood effects. The paper provides a review of recent advances in the program evaluation literature in order to make a clear distinction between the interpretation of Intent to Treat (ITT) and Treatment on the Treated (TOT) parameters as program effects and Local Average Treatment Effect (LATE) parameters as neighborhood effects. Due to the nature of the LATE, this distinction helps to clarify that results from MTO are only informative about a small

[1] Although this number did drop in the 1990s (Jargowsky (2003)), it was not nearly enough to return to 1970 levels (Aliprantis and Zenker (2011)).

[2] The neighborhood effects considered in this paper are those associated with living in a neighborhood in the US characterized by this "new urban poverty" (Wilson (1996)), but there are many alternative definitions of neighborhood effects (Durlauf (2004)).

subset of neighborhood effects of interest. The evaluation framework is also used to emphasize that since the LATE is defined by the subgroup of compliers, different instruments will result in different LATE parameters if they induce different subpopulations to select into treatment (Heckman (1997)). An important implication is that alternative housing mobility programs designed to induce moves to neighborhoods with characteristics in addition to or in lieu of low poverty might induce larger effects than MTO.

After this review of the literature, the paper uses experimental group status in MTO as an instrumental variable to estimate the LATEs of various neighborhood characteristics. A first step in this process is to determine at which sites experimental group status was a strong instrument for neighborhood treatments. Tests for instrument strength show that MTO induced large changes in neighborhood poverty rates. However, it is also shown that MTO induced remarkably little variation in many of the other neighborhood and school characteristics believed to influence outcomes and that much of this variation was confined to the tails of these characteristics' national distributions. Some of the neighborhood characteristics that can be described in this way include school quality, the female high school graduation rate, and the share of single-headed households in participants' neighborhood of residence. This step also shows substantial differences in outcomes across MTO sites, pointing to the investigation of heterogeneity in program effects from MTO as a fruitful direction for future research.

LATE estimates at sites where experimental group status was a strong instrument are consistent with prevailing theories of neighborhood effects. Moves to neighborhoods with low poverty rates, a high degree of personal safety, or a high female labor force participation rate are all associated with increases in labor force participation. Moves to low poverty and safe neighborhoods are also associated with improved health outcomes. And although improvements in labor market outcomes such as employment and income coming from moves to such neighborhoods are not estimated precisely enough to be statistically significant, these effects are improvements and they are of large magnitudes.

The paper proceeds as follows: Section 2 describes the MTO experiment and presents some descriptive statistics. Section 3 draws heavily from a line of research by Heckman and coauthors (See Heckman (2010), Heckman et al. (2006), and Heckman and Vytlacil (2005).) to define and interpret several treatment effect parameters when assigned treatment is viewed as an instrumental variable. A summary of the program effects found in the literature is also presented in this section. Section 4 discusses the data used in the analysis. Section 5 uses these data to identify the sites at which the experiment is strongly associated with participants selecting into neighborhoods with various characteristics of interest. Section 5 then presents a discussion of the assumptions used in estimation, followed by the main estimation results. Section 6 discusses the implications of these results for our understanding of neighborhood effects and the design of housing mobility programs. Section 7 concludes.

2 Moving To Opportunity (MTO)

Moving To Opportunity (MTO) was inspired by the promising results of the Gautreaux program. Following a class-action lawsuit led by Dorothy Gautreaux, in 1976 the Supreme Court ordered the Department of Housing and Urban Development (HUD) and the Chicago Housing Authority (CHA) to remedy the extreme racial segregation experienced by public-housing residents in Chicago. One of the resulting programs created by HUD and CHA gave families awarded Section 8 public housing vouchers the ability to use them beyond the territory of CHA. The results from the Gautreaux program came at the same time that much attention was being devoted to the increasing concentration of poverty in the US (Wilson (1987)), and they indicated that housing mobility could be an effective policy to mitigate the adverse effects of segregation and concentrated poverty.

The Gautreaux court ruling allowed families to be relocated either to suburbs that were less than 30 percent black or to black neighborhoods in the city that were forecast to undergo "revitalization" (Polikoff (2006)). Although families awarded Section 8 certificates were eventually trained to find their own housing, the initial relocation process of the Gautreaux program created a quasi-experiment, as families at the top of a waiting list were matched to neighborhoods based on the availability of housing units (Polikoff (2006)). Relative to city movers, suburban movers from Gautreaux were more likely to be employed (Mendenhall et al. (2006)), and the children of suburban movers attended better schools, were more likely to complete high school, attend college, be employed, and had higher wages than their city mover counterparts (Rosenbaum (1995)).[3]

In the wake of this promising evidence from Gautreaux, there was bipartisan support for attempts to deconcentrate poverty and improve outcomes through housing vouchers (Goering (2003)). The Housing and Community Development Act of 1992 authorized HUD to "assist very low-income families with children who reside in public housing or housing receiving project-based assistance ... to move out of areas with high concentrations of persons living in poverty to areas with low concentrations of such persons" (Goering (2003)).[4] MTO offered housing vouchers to eligible households between September 1994 and July 1998 in five US cities; Baltimore, Boston, Chicago, Los Angeles, and New York (Goering (2003)). Households were eligible to participate in MTO if they were low-income, had at least one child under 18, were residing in either public housing or Section 8 project-based housing located in a census tract with a poverty rate of at least 40%, were current in their rent payment, and all families members were on the current lease and were without criminal records (Orr et al. (2003)).

In addition to implementing the program, Congress also required that HUD conduct evaluations of the demonstration (Goering (2003)). HUD contracted with Abt Associates to implement a social experiment by randomly assigning households to various treatments. This was achieved by adding

[3]It has also been found that suburban movers have much lower male youth mortality rates Votruba and Kling (2009) and tend to stay in high-income suburban neighborhoods many years after their initial placement (DeLuca and Rosenbaum (2003), Keels et al. (2005)).

[4]The threshold for high-poverty was set to follow a common cutoff considered in the social sciences, census tracts where 40% or more of residents are poor (Jargowsky (1997)), while the threshold of low poverty was set at the median tract-level poverty rate in 1990, 10% (Goering (2003)).

households to a waiting list after they volunteered to take part in MTO. Between 1994 and 1997 families were drawn from the waiting list through a random lottery. After being drawn, families were randomly allocated into one of three treatment groups; the *experimental* group, the *Section-8 only* comparison group, and the *control* group. The *experimental* group was offered Section 8 housing vouchers, but were restricted to using them in census tracts with 1990 poverty rates of less than 10 percent. However, after one year had passed, families in the *experimental* group were then unrestricted in where they used their Section 8 vouchers. Families in this group were also provided with counseling and education through a local non-profit. Families in the *Section-8 only* comparison group were provided with no counseling, and were offered Section 8 housing vouchers without any restriction on their place of use. And families in the *control* group received project-based assistance.

Out of 4,610 families that applied, there were 4,248 accepted families who participated in MTO, with 1,310 families in the control group, 1,209 families in the Section 8 only group, and 1,729 families in the experimental group (Sanbonmatsu et al. (2006), Clark (2008)). Around two-thirds of the families who volunteered for the program were African-American, while most of the rest were Hispanic (Kling et al. (2005), Table F13 in Kling et al. (2007b)). About 25 percent of eligible families applied to participate in MTO (Ludwig et al. (2001)). Compared to those who did not move, those in the treatment groups who moved through MTO were younger, more likely to have no teenage children, to have reported a neighborhood that is very unsafe at night, to have been very dissatisfied with their apartment, to have been enrolled in school, and to have had confidence in their ability to move through the voucher program (Kling et al. (2007a)). More information on MTO may be found on HUD's MTO webpage or the NBER online repository of papers on MTO.

3 The Identification of Treatment Effects in Social Experiments

In order to think about neighborhood effects from MTO, we now consider a standard framework for studying causal treatment effects (Holland (1986), Rubin (1974), Heckman and Vytlacil (2005)). Let $Y_i(1)$ and $Y_i(0)$ be random variables associated with the potential outcomes in the treated and untreated states, respectively, for individual i. D_i is a random variable indicating receipt of a binary treatment, where

$$D_i = \begin{cases} 1 & \text{if treatment is received;} \\ 0 & \text{if treatment is not received.} \end{cases}$$

The measured outcome variable Y_i is

$$Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0). \tag{1}$$

Since both treatment states are not observable for any individual i, inference cannot be drawn about the value of $Y_i(1) - Y_i(0)$. However, causal inference about population averages can be made under specific assumptions. One such assumption that allows for inference about average effects on a population, which Holland (1986) calls **Independence**, is that:

$$E[Y_i(1)] = E[Y_i(1)|D_i = 1]$$
$$E[Y_i(0)] = E[Y_i(0)|D_i = 0].$$

This assumption is typically operationalized by the researcher's random assignment of individuals to treatment. When true, this assumption yields

$$\frac{\sum_{i=1}^{I} D_i Y_i}{\sum_{i=1}^{I} D_i} - \frac{\sum_{i=1}^{I}(1-D_i)Y_i}{\sum_{i=1}^{I}(1-D_i)}$$

as an unbiased estimator of the Average Treatment Effect (ATE):[5]

$$\beta^{ATE} = E[Y(1) - Y(0)].$$

There are two reasons the ATE defined above is typically not the primary interest of researchers in the social sciences. First, the ATE measures the average response to treatment in the entire population. Nearly all social programs are targeted to a specific subpopulation hypothesized to benefit from the program. Second, it is rarely feasible to estimate the ATE in social settings. Individuals are able to choose whether or not to participate in programs, such as job training programs (LaLonde (1995)), Head Start (Ludwig and Miller (2007), Garces et al. (2002)), or housing mobility programs like Gautreaux and MTO.

3.1 Identification of Program Effects

3.1.1 The ITT

In the case of social experiments, a researcher can typically control assignment but not receipt of treatment. Thus we define Z as an indicator for the treatment assigned to an individual:

$$Z = \begin{cases} 1 & \text{if treatment is assigned;} \\ 0 & \text{if treatment is not assigned.} \end{cases}$$

If we assume

Assumption 1* $Y(0)$ and $Y(1)$ are jointly independent of Z

Assumption 2* $E[Y(0)] < \infty$ and $E[Y(1)] < \infty$,

[5] From this point forward individual subscripts i will be dropped, but it is understood that expectations are taken over the population of individuals.

then by comparing the outcome variable Y at two different values of assigned treatment, $Z = 1$ and $Z = 0$, we obtain the Wald estimator:

$$
\begin{aligned}
&E[Y|Z=1] - E[Y|Z=0] \\
&\qquad = E[D(1)Y(1) + (1-D(1))Y(0)|Z=1] - E[D(0)Y(1) + (1-D(0))Y(0)|Z=0] \\
&\qquad = E[(D(1)-D(0))(Y(1)-Y(0))] \qquad (2)\\
&\qquad = Pr[D(1)-D(0)=1]\,E[Y(1)-Y(0)|D(1)-D(0)=1] \qquad (3)\\
&\qquad\quad + Pr[D(1)-D(0)=-1]\,E[Y(0)-Y(1)|D(1)-D(0)=-1].
\end{aligned}
$$

Equation 2 follows from Assumption 1^*, and Assumption 2^* ensures the Wald estimator is finite.

One causal parameter of interest is the Intent-to-Treat (ITT) effect, which is the causal effect of treatment assignment on outcomes:

$$\beta^{ITT} \equiv E[Y|Z=1] - E[Y|Z=0]. \tag{4}$$

The only assumptions necessary to identify the ITT effect from the Wald estimator in Equation 3 are Assumptions 1^* and 2^*. Under these assumptions, Equation 3 represents a comparison of weighted average outcomes between those individuals who "switch-in" (compliers) and those who "switch-out" (defiers) of treatment due to changes in assigned treatment. The outcomes of those whose treatment is not affected by assigned treatment, always-takers and never-takers, do not contribute to this estimate.

3.1.2 A Brief Review of Program Effect Estimates from MTO

The Intent-to-Treat (ITT) effects identified in the literature on MTO compare the mean outcomes of those offered a housing voucher with the outcomes of households who were not offered a housing voucher. These parameters have a clear policy interpretation: they are the effects on outcomes from being offered a housing voucher through the MTO program. And since the offer of a housing voucher ($Z = 1$, or assigned treatment) was randomly allocated to households, these effects should be interpreted as causal effects. Based on the outcomes of Gautreaux, researchers expected to find universally positive effects of moving through MTO (Kling et al. (2007a), Sanbonmatsu et al. (2006)). In contrast to researchers' expectations, the data show that the effects of the program were mixed.

There were no significant effects on earnings, welfare participation, or the amount of government assistance adults received 5 years after randomization (Kling et al. (2007a)). There was also little effect on adult physical health: No statistically significant effect on self-reported overall health, hypertension, or asthma (Kling et al. (2007a)). The single improvement in adult outcomes - a 5 percentage point reduction in adult obesity for the experimental group relative to the control group - cannot be distinguished from statistical aberration since there are multiple hypotheses being tested simultaneously (Kling et al. (2007a)).

However, there were positive ITT effects on measures of adult mental health such as distress and calmness (Tables III in Kling et al. (2007a) and F5 in Kling et al. (2007b)). In fact, the magnitude of the improvements in adult mental health were comparable to the most effective clinical and pharmacological interventions (Kling et al. (2007a)). Kling et al. (2007a) hypothesize that this improvement in mental health is due to a reduction in the fear of random violence. A related outcome is that adults in the experimental group were much less likely to report that police do not come when called in the neighborhood (Table II and p 102 of Kling et al. (2005)).

Improved outcomes for young females were found in the groups offered a housing voucher through MTO. For young females ages 15-25 in 2001 (4-7 years after randomization), Kling et al. (2005) find that the effect of being assigned to the experimental group is about one-third fewer arrests for violent and property crimes relative to the control group. Kling et al. (2007a) analyze results from MTO youth aged 15-20 at all five sites an average of five years after random assignment. They find positive ITT effects for female youth that are largest with respect to mental health and still substantial for education and risky behavior (Kling et al. (2007a), Table G2 in Kling et al. (2007b)).

MTO had negative ITT effects on the outcomes of young males. The effects on young males were a deterioration in physical health and an increase in risky behavior, smoking and non-sports injuries (Kling et al. (2007a)), as well as an increase in the fraction of days absent from school and the probability of having a friend who uses drugs (Kling et al. (2005), Table IX). While Kling et al. (2005) find statistically insignificant changes in violent crime arrests, they also find a positive ITT effect of about one-third of the control group mean for property crime arrests. After considering empirical evidence on three reasons for these gender differences - peer sorting, coping strategies, and a comparative advantage in property offending - Kling et al. (2005) conclude that these outcomes result from boys being more likely to take advantage of a newfound comparative advantage in property offending in their new neighborhoods.[6] The dynamics of these behaviors are interesting, as young males have significantly lower violent crime arrests in the first two years after random assignment, but property crime rates then increase significantly starting 3 and 4 years after assignment (Kling et al. (2005), Table V).

Sanbonmatsu et al. (2006) examine test score data collected in 2002 for MTO children who were 6-20 on December 31, 2001 and find no evidence of improvements in reading scores, math scores, behavior problems, or school engagement. Sanbonmatsu et al. (2006) first combine reading and math test scores (Woodcock-Johnson Revised scores) and estimate ITT effects for all ages, as well as by subgroups of 6-10, 11-14, and 15-20. The ITT effects for the combined reading and math scores are neither statistically significant for any age subgroup nor for all ages together (p 673). When Sanbonmatsu et al. (2006) examine ITT effects for several other educational outcomes, such as grade repetition, suspensions, measures of school engagement such tardiness and paying

[6] An interesting note from Kling et al. (2005) is that these effects seem to be driven by the number of arrests for those who are criminally involved, rather than the rate of participation in criminal activity (p 102). However, effects are similar for those with and without histories of anti-social behaviors prior to random assignment, such as arrest, expulsion from school, or parents called to school for problems (p 112).

attention in class, they find only one ITT effect to be statistically significant: the effect of being offered a voucher actually increases problem behaviors for youth aged 11-14 (p 673).

The only achievement test effect for subgroups that is statistically significant is a positive experimental ITT on reading for African-American children (p 678). The positive impacts on test scores were only found in Baltimore and Chicago (Sanbonmatsu et al. (2006), p 678; Burdick-Will et al. (2010)). These sites were almost entirely African-American (unlike the other sites, which had many Hispanic households), and also had higher crime rates (Burdick-Will et al. (2010)).

3.2 Identification of Neighborhood Effects

3.2.1 Assigned Treatment and Selection into Treatment

In addition to using the results from MTO for studying the effects of housing voucher programs, researchers have also interpreted estimates of the effects of moving through MTO as neighborhood effects. Since several of the parameters in the program evaluation literature relevant to this research are defined in terms of the subpopulation receiving treatment, we consider a model of how individuals select into treatment. We begin by noting that it need not be true that $D = Z$, and so we write $D(Z)$ to denote the treatment received when assigned treatment Z. We next suppose there is a latent index D^* that depends on assigned treatment Z and some unobserved component U^* as follows:

$$D^* = \mu^*(Z) - U^*, \tag{5}$$

and that individuals select into treatment status based on their latent index:

$$D = \begin{cases} 1 & \text{if } D^* \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

We follow Heckman and Vytlacil (2000) and assume:

Assumption 1 $Y(0)$, $Y(1)$, and $D(z)$ are jointly independent of Z

Assumption 2 $E[Y(0)] < \infty$ and $E[Y(1)] < \infty$

Assumption 3 $\mu^*(Z)$ is a non-degenerate random variable

Assumption 4 $U^* \sim U[0, 1]$

Note that there is no loss of generality for the selection model in Equations 5 and 6 by making Assumptions 3 and 4. As noted in Heckman and Vytlacil (2000), Assumptions 3 and 4 imply that if $D^* = v(Z) - V$, we may equate the two models by writing $\mu^*(Z) = F_V(v(Z))$ and $U^* = F_V(V(Z))$.[7] We write the propensity score as $P(z) = Pr(D = 1|Z = z)$. Note that Assumption 4

[7] Applying the probability integral transformation is much less useful in the variable treatment intensity case relative to the binary case presented here.

implies $\mu^*(z) = P(z)$ when $\mu^*(z) \in [0,1]$. In the discussion that follows, $\mu^*(z)$ and $P(z)$ are used interchangeably depending on which term better facilitates interpretation.

Table 1 shows how the labels in Angrist et al. (1996) apply to individuals due to their response to treatment assignment. Figure 2 shows how these labels are generated by the selection model in Equations 5 and 6, and we focus on the case displayed in Figure 2a. Note that since the unobservable component of the latent index is distributed according to a uniform $[0,1]$ distribution, $U^* \sim U[0,1]$, treatment does not depend on U^* or Z if both $\mu^*(Z=0) < 0$ and $\mu^*(Z=1) < 0$, or if both $\mu^*(Z=0) \geq 1$ and $\mu^*(Z=1) \geq 1$. Specifically, if $\mu^*(Z) < 0$ for both $Z = 0, 1$, then an individual is a never-taker, while if $\mu^*(Z) \geq 1$ for both $Z = 0, 1$, then an individual is an always-taker.

It is when $\mu^*(Z) \in (0,1)$ that treatment depends on both assigned treatment and the unobserved component of the latent index, U^*. Consider the situation portrayed by $D^* = \mu_1^*(Z) - U_1^*$ in Figure 2a. In the case that $0 < \mu_1^*(Z=0) < \mu_1^*(Z=1) < 1$, assigning treatment to an individual makes them more likely to participate. Individuals with $u^* \in [0, \mu_1^*(Z=0))$ are always takers, those with $u^* \in [\mu_1^*(Z=0), \mu_1^*(Z=1))$ are compliers, and those with $u^* \in [\mu_1^*(Z=1), 1)$ are never-takers.[8]

Note that if $D^* = \mu_1^*(Z) - U_1^*$ for all individuals as in Figure 2a, then there are no defiers.[9] Furthermore, if all individuals select into treatment according to $D^* = \mu_1^*(Z) - U_1^*$ and it is the case that $\mu_1^*(Z=0) \leq 0$ and $\mu_1^*(Z=1) \in (0,1)$, then there are no always takers, only compliers and never-takers. When $D = 1$ is defined as use of a voucher offered by the MTO program, this is a reasonable way of modeling selection into treatment, as families could not have used a voucher through the MTO program unless they were assigned a voucher through the MTO program. However, under alternative definitions of $D = 1$, particularly those in which $D = 1$ is moving to a neighborhood with a particular characteristic, assuming $\mu_1^*(Z=0) \leq 0$ may be unreasonable. Being able to say whether $\mu_i^*(Z=0) < 0$ or $\mu_i^*(Z=0) \in (0,1)$ will depend on the definition of treatment, which will in turn determine how we interpret parameter estimates.

Given our joint model of outcomes (Equation 1) and selection into treatment (Equations 5 and 6), we now consider the assumptions necessary to identify parameters of interest. We will use this joint model to define and interpret these parameters.

3.2.2 The TOT and LATE

Researchers are often interested in how *receiving* treatment affects outcomes. Since the ITT parameter can only be interpreted as the effect of *assigning* treatment to units/individuals, it is uninformative to researchers on this topic. Thus there is interest in using Equation 3 to identify

[8]Throughout this paper the word complier will be defined as in Angrist et al. (1996).

[9]In order for the selection model in Equations 5 and 6 to produce defiers, some subpopulation would have to select into treatment according to $D^* = \mu_1^*(Z) - U_1^*$ with $\mu_1^*(Z=1) > \mu_1^*(Z=0)$, while another subpopulation would have to select into treatment according to $D^* = \mu_2^*(Z) - U_2^*$ with $\mu_2^*(Z=1) < \mu_2^*(Z=0)$. This could be an example of essential heterogeneity as defined in Heckman et al. (2006). Related examples include the way parents select their children into the treatment of kindergarten entrance age (Aliprantis (2011)) and the way students select into attaining a GED, graduating from high school, or dropping out of high school in response to easing GED requirements (Heckman and Urzúa (2010)). The assumption of monotonicity introduced in Imbens and Angrist (1994), and presented shortly in Assumption 5, rules out the possibility of defiers.

treatment effects that go beyond the ITT and inform us about the effect of treatment on outcomes.[10] Much of the literature on instrumental variables does this by placing restrictions on how changes in the instrument induce changes in treatment (ie, on the selection model in Equations 5 and 6.).

In the case of a social experiment like MTO, the instrument is assigned treatment (Heckman (1996)). Imbens and Angrist (1994) and Angrist and Imbens (1995) develop several assumptions made on the selection model in Equations 5 and 6 that allow for the identification of treatment effects when combined with Assumptions 1–4. In the context of our selection model, the monotonicity assumption introduced in Imbens and Angrist (1994) is:

Assumption 5a $\mu^*(Z=0) < \mu^*(Z=1)$ for all individuals

Assumption 5b At least one of $\{\mu^*(Z=0), \mu^*(Z=1)\}$ is in $(0,1)$ for all individuals

Assumption 3 implies that $\mu^*(Z=0) \neq \mu^*(Z=1)$, and Assumption 5 ensures that being assigned to treatment makes no individuals less likely to receive treatment, while at the same time ensuring that some individuals are induced to receive treatment due to the instrument. That is, together with Assumptions 1-4, Assumption 5 ensures that $Pr[D(1)-D(0)=-1]=0$ and $Pr[D(1)-D(0)=1] \neq 0$, so the Wald estimator from Equation 3 identifies

$$\beta^{LATE}(Z=1, Z=0) \equiv E[Y(1)-Y(0)|D(1)-D(0)=1] = \frac{E[Y|Z=1]-E[Y|Z=0]}{Pr[D(1)-D(0)=1]}. \tag{7}$$

The Local Average Treatment Effect (LATE) is the average effect of treatment on outcome Y for those who can be induced to change treatment by a change in assigned treatment (ie, The LATE informs us about the average effect of treatment on compliers.).

The Treatment-on-the-Treated (TOT) effect is defined as the average change in outcome for those who are treated, or the average effect of treatment over both compliers and always-takers. It is possible for the researcher to identify this parameter if they believe there are no always-takers (ie, that $Pr[D(0)=1]=0$). In the context of our selection model, we might assume:

Assumption 6 $\mu^*(Z=0) < 0$ and $\mu^*(Z=1) \in (0,1)$ for all individuals

Under Assumptions 1-4 and 6, the Wald estimator allows us to identify

$$\beta^{TOT} \equiv E[Y(1)-Y(0)|D(1)=1] = \frac{E[Y|Z=1]-E[Y|Z=0]}{Pr[D(1)=1]}. \tag{8}$$

Remember that within our selection model the subgroups of compliers, always-takers, and never-takers can be identified by the interval in which their realization of U^* lies. Thus, as pointed out in Heckman and Vytlacil (2000), the TOT and LATE parameters can be seen as the average value of $Y(1)-Y(0)$ for U^* lying in different intervals. Under the selection model in Equations 5 and 6,

[10] To be clear, this paper refers to the treatment actually received by a unit as *treatment* and the treatment assigned to a unit as *assigned treatment*.

together with Assumptions 1-5, we have:

$$\beta^{TOT} = E[Y(1) - Y(0) \mid 0 \leq U^* \leq \mu^*(Z=1)]$$
$$\beta^{LATE}(Z=1, Z=0) = E[Y(1) - Y(0) \mid \mu^*(Z=0) \leq U^* < \mu^*(Z=1)].$$

That is, β^{TOT} is the average response to treatment for individuals with $U^* \in (0, \mu^*(Z=1))$, while β^{LATE} is the average response to treatment for individuals with $U^* \in (\mu^*(Z=0), \mu^*(Z=1))$. Since $U^* \in [0,1]$, Assumption 6 implies there will be no always-takers and thus the LATE and TOT parameters will be identical:

$$\begin{aligned}\beta^{LATE}(Z=1, Z=0) &= E[Y(1) - Y(0) \mid \mu^*(Z=0) < 0 \leq U^* < \mu^*(Z=1)] \\ &= \beta^{TOT}.\end{aligned}$$

3.2.3 The MTE and the Interpretation of the TOT and LATE

Heckman and Vytlacil (2005) show that the TOT and LATE, in addition to several other treatment effects, may be defined in terms of the Marginal Treatment Effect (MTE). We consider some of their results here to clarify the interpretation of TOT and LATE estimates. Given the latent index model in Equation 6, we define the MTE as

$$\beta^{MTE}(u) = E[Y(1) - Y(0) | U^* = u].$$

Heckman et al. (2006) show that while Assumptions 1-4 are met and $Pr(D=1) \in (0,1)$ we can interpret the MTE as the derivative of the expected value of the outcome at a particular level of U^*, so that:

$$\begin{aligned}E[Y|P(Z) = u''] &= E[Y|P(Z) = u'] + \int_{u'}^{u''} \frac{\partial}{\partial p} E[Y|P(Z) = x] dx \\ &= E[Y|P(Z) = u'] + \int_{u'}^{u''} E[Y(1) - Y(0)|P(Z) = x] dx.\end{aligned}$$

Furthermore, the Mean Value Theorem tells us there exists some $u^\star \in (u', u'')$ such that:

$$\frac{E[Y|P(Z) = u''] - E[Y|P(Z) = u']}{u'' - u'} = E[Y(1) - Y(0)|P(Z) = u^\star].$$

If we let $P(Z=1) = u''$ and $P(Z=0) = u'$, then under Assumptions 1-5 we can think of the LATE as the MTE evaluated at that individual for which the MTE is equal to the average MTE over the interval in question. The TOT is the limiting case of the LATE parameter as $P(0) = u' \downarrow 0$.

$\beta^{MTE}(u)$ is the mean response to treatment for persons with $U^* = u$, and the previous discussion illustrates that the LATE is the average value of $\beta^{MTE}(u)$ for persons with u lying in some interval (u', u''). An important point is that the interval over which the LATE is identified is determined by the values of $\mu(Z)$, with $(u', u'') = (P(Z=0), P(Z=1))$. Further discussion of these parameters may be found in Heckman and Vytlacil (2000), Heckman and Vytlacil (2005), Heckman (2010), Angrist et al. (1996), and Imbens and Angrist (1994).

3.2.4 Two Hypothetical Examples in the Context of MTO

Now consider a hypothetical example from MTO in which the outcome of interest Y is hours worked per week. Figure 3 illustrates this example under two possible definitions of treatment to clarify the interpretation of the treatment effects just discussed. Look first at the treatment effects in Example 1. This example shows that because the LATE allows for heterogeneity in response to treatment in a general way, the LATEs for individuals with u in different intervals can be dramatically different. We see that while $\beta^{LATE}(0.1, 0.6)$ and $\beta^{LATE}(0.6, 0.9)$ are of similar magnitudes in this example, they are of opposite signs. And although $E[Y|P(Z) = 0.3]$ is approximately the same as $E[Y|P(Z) = 0.6]$, the magnitudes of $\beta^{LATE}(0.1, 0.3)$ and $\beta^{LATE}(0.1, 0.6)$ are quite different because the second LATE is an average over a much larger share of individuals. In this example treatment has a large positive effect on hours worked per week on those individuals most likely to select into treatment, those with small values of u, and this effect decreases to the point of becoming negative for those difficult to induce into treatment, those with large values of u.

Now consider Example 2. Since the LATE framework allows for general heterogeneity in response to treatment, we need not know a priori whether the function $E[Y|P(Z) = p]$ looks like it does in Example 1 or Example 2. It could be the case that $\beta^{LATE}(0.1, 0.3) = 45$ and $\beta^{LATE}(0.6, 0.9) = -24$ as in Example 1, or it could be the case that $\beta^{LATE}(0.1, 0.3) = 1$ and $\beta^{LATE}(0.6, 0.9) = 42$ as in Example 2. Knowledge of one LATE for individuals with U^* in a given interval need not be informative about the LATE for individuals in any other interval. Furthermore, it is highly unlikely that we get to choose the interval of U^* realizations to which a LATE pertains. In the case of MTO, the intervals are determined by the share of households selecting into treatment in both the experimental and control groups. Example 2 also helps to highlight the difference between the TOT and LATE. Suppose there is a large share of always-takers, all those with $u \in (0, 0.6)$. Then the average effect on compliers (the LATE) could be dramatically larger than the average effect on always-taker and compliers together (the TOT). This is because in Example 2 treatment only has a small effects on hours worked for those with low values of u, those most likely to select into treatment, but has effects that become quite large for those with high values of u, those least likely to select into treatment.

These examples help to show how important selection into treatment is for interpreting LATE estimates. A given instrument such as experimental group status in MTO only identifies one LATE for a given definition of treatment. A complete understanding of the effects of a given treatment in our framework would require knowledge of the MTE at all values of U^*, yet a given LATE only informs us about the average MTE for individuals in a specific interval (u', u''), which is itself determined by how households select into treatment. In the case of MTO, LATE parameters are the effects of moves to neighborhoods with particular characteristics *induced by MTO.* Other instruments would likely have induced other subpopulations to select into the defined neighborhood treatment, and thus would almost certainly yield different LATEs for a given definition of treatment. For example, although only about 20% of families participating in Gautreaux moved through the program, its alternative rules for moving likely induced a greater share of compliers for many neighborhood treatments than did MTO, despite the fact that approximately 47% of experimental families moved through the program (Ludwig et al. (2008)). Alternative housing mobility programs designed to induce moves to different types of neighborhoods could plausibly result in different neighborhood effects. Thus while we may estimate neighborhood effects from MTO, even an experiment such as MTO can only inform us about a small subset of the neighborhood effects of interest within our framework.

3.3 Program Effects Versus Neighborhood Effects from MTO

The interpretation of results from MTO as neighborhood effects has created controversy among social scientists. One interpretation argues that selection into treatment biases parameter estimates (Clampet-Lundquist and Massey (2008)). An opposing interpretation is that randomization makes selection unimportant for the estimation and interpretation of treatment effects from MTO. A quote from Ludwig et al. (2008) summarizes this view: "Randomization ... solves the selection problem, by causing variation in neighborhood of residence to occur for reasons that are uncorrelated with individual characteristics, whether or not those characteristics are measurable."[11]

The preceding discussion of the LATE helps to reconcile these opposing views. Note that "The critical feature of the problem of evaluating a treatment under imperfect compliance is that even if assignment Z_i is random or ignorable, the actual receipt of treatment D_i is typically nonignorable [or nonrandom]" (Angrist et al. (1996), p 447). Therefore the selection model in Section 3 is appropriate for evaluating results from MTO. Randomization occurred in MTO at the level of assigned treatment ($Z \in \{0, 1\}$), not at the level of treatment ($D \in \{0, 1\}$). Households were able to choose whether or not to move, and where to move to, after receiving their assigned treatment.

While it is true that randomization in MTO did indeed induce variation in neighborhood of residence that was uncorrelated with individual characteristics, our model of selection helps us to see that this statement is only true *within* the subgroup of compliers. This point does not impact the current interpretation of program effect estimates found in the literature. One set of such effects are ITT effects for which treatment is defined as "being offered a housing voucher through MTO." Another set of program effects are TOT parameters for which treatment is defined as "moving through the MTO program." Since no one can move through the MTO program unless assigned treatment, there are no always-takers of this treatment, and thus the TOT and LATE parameters are identical. These TOT effects are all qualitatively similar to the ITT effects reviewed earlier.[12]

[11]This view is shared by most of the influential articles on MTO, including Kling et al. (2007a), Kling et al. (2005), Sanbonmatsu et al. (2006), Ludwig (2010), and even Sampson (2008).

[12]The TOT effects have a slightly larger magnitude since they are the ITT effects divided by the probability of receiving treatment, which in this case is estimated in Table F9 of Kling et al. (2007b) to be 0.467. For example, the TOT effect of this treatment was a 10 percentage point decrease in adult obesity, compared with an ITT effect of a 5 percentage point decrease (Kling et al. (2007a)).

On the other hand, exogenous variation in neighborhood of residence being restricted to the subgroup of compliers will have major implications if we are interested in learning about the effects of specific neighborhood characteristics on outcomes. If treatment were defined to be "moving to a neighborhood with characteristic x," where x were low poverty, high employment, or a high degree of personal safety, then there would be always-takers, and the TOT and LATE parameters would no longer be the same. As discussed in Section 3.2.4, LATE estimates may not only differ dramatically in magnitude from ITT or TOT estimates, but their interpretation may also be quite different from that of the ITT and TOT estimates found in the literature. Using LATEs to assess the evidence on neighborhood effects from MTO will likely lead to different conclusions than using ITT or TOT effects for the same purpose.

4 Data

The primary source of data is the MTO Interim Evaluation sample. In addition to information on the MTO sample collected both at the time members volunteered to participate and during MTO participation, the Interim Evaluation data also contain information on the MTO sample in 2002, the time the evaluation was conducted. Each individual was linked to their household in two ways, and these definitions are necessary for defining the two types of individuals in this sample whose outcomes we will investigate. The base household members were those individuals living together at the baseline, or the period before random assignment, and the core household members were those individuals planning to move together if awarded a voucher. For adult outcomes we consider the outcomes of adult females who are core household members. For youth outcomes we consider children between the ages of 5 and 19 as of May 31, 2001. Up to two such children from each core household were included in the sample, and in households with more than two children, two children were randomly selected.

The MTO Interim Evaluation data contain some variables about the census tracts in which participants resided that will be used to understand neighborhood characteristics. A secondary source of data on neighborhood characteristics used in the analysis is decennial census data from the National Historical Geographic Information System (NHGIS, Minnesotra Population Center (2004)). Tract-level variables from these data are matched to the MTO Interim Evaluation sample to supplement the variables related to neighborhood characteristics that are available in the Interim Evaluation data.

4.1 Variables

Some treatments are defined in terms of neighborhood characteristics, which are measured at the level of census tracts. When available, tract-level variables from the MTO Interim Evaluation are used before using the matched tract-level variables from the NHGIS census data. The Interim Evaluation has variables reporting the 2000 census data for the census tracts in which core households lived in 2002. These variables used in the analysis include poverty rate and percent minority. The sample also lists the census tract in which core households lived in 2002, and this information is linked to NHGIS variables from the 2000 census. The variables created in this way include the female high school graduation rate, the female BA attainment rate, the female labor force participation rate, the percent of females employed in a management, professional, or related occupation, the percent of households receiving public assistance income in 1999, and the share of households with own-children under the age of 18 that are single-headed.

Supplementing these census data are variables obtained by directly asking respondents about their neighborhood of residence in 2002. These variables include whether the respondent feels safe in the current neighborhood at night, whether there is a problem with police not responding when called to the neighborhood, and whether the respondent believes it is likely or very likely that neighbors would do something about children spray-painting graffiti or skipping school.

In addition to these neighborhood characteristics, many children's treatments are defined in terms of schools or peer groups. One set of school variables are constructed as weighted averages of characteristics of all schools attended by the child using administrative data; these include the percentile rank on state assessments and the student-teacher ratio. Another set of school variables are children's responses to questions about the most recent school they have attended. These questions include whether the child agrees or strongly agrees that they feel safe in school, whether the child agrees or strongly agrees that teachers are interested in students, whether the school has discipline or disruption problems, and how the child rates the overall school climate using five subquestions. There are also variables on the child's peer group, which include whether they have any friends who use drugs or carry a weapon, and whether they live in the same neighborhood as at the baseline or visit with friends from that neighborhood.

The labor market outcomes of adults are variables measuring the self-reported total earnings of the head of household, the total household income (all sources summed), the individual earnings in 2001 of the sample adult, the labor market status of the adult at the time of the interim survey (ie, Whether they are employed, unemployed, or not in the labor force at all.), and a binary variable indicating whether an adult respondent or her children have received welfare benefits (AFDC/TANF) at any time during the past two years. Adult health outcomes used are a psychological distress index for the adult respondent that is the fraction of six psychological distress items that the adult reported feeling at least some of the time during the past month, the adult's Body Mass Index (BMI), and whether the adult reported symptoms of depression during the past year.

Youth outcomes used include problem behaviors, such as whether a child has ever smoked a cigarette, ever smoked pot, ever drunk alcohol, or ever been arrested. The measure of education and labor market outcomes is a binary variable indicating whether a child was idle during the past week, where idle is defined as neither being in school nor being employed. Health outcomes of youth include whether the child reports having an asthma attack or wheezing in the past year, the BMI percentile of the youth, whether the youth has ever had depression symptoms, and a psychological distress index for children that is the fraction of the six psychological distress items the child reported feeling at least some of the time during the 30 days prior to the interim survey. Test score data are not used in examining youth outcomes for three reasons. First, careful analysis has already been done of these data (Sanbonmatsu et al. (2006), Burdick-Will et al. (2010)). Second, there are some surprising patterns in these data that are difficult to explain, such as the fact that MTO children scored higher on tests than one would predict from their demographic characteristics (See Sanbonmatsu et al. (2006), p 659.). And most importantly, MTO only induced small changes in school quality. This variation will be examined closely in Section 5.2.

4.2 Weights

Weights are used in all estimates for two reasons. First, random assignment ratios varied both from site to site and over different time periods of sample recruitment. Randomization ratio weights are used to create samples representing the same number of people across groups within each site-period. This ensures neighborhood effects are not conflated with time trends. Second, sampling weights must be used to account for the sub-sampling procedures used during the interim evaluation data collection.

4.3 Descriptive Statistics at Baseline

The MTO Interim Evaluation sample used in this analysis includes 4,156 adult females and 6,683 children. Table 2 shows some descriptive statistics of this adult sample at the baseline. At baseline 38% of adults reported having completed high school, and an additional 19% reported having completed a GED. 25% of adult respondents were working for pay, and 75% were receiving AFDC/TANF benefits. 50% reported that the streets near their home were very unsafe at night, and 42% report that during the 6 months preceding the survey, a household member had been beaten/assaulted; threatened with a gun or knife; or had their purse, wallet, or jewelry snatched from them. Only 16% of adult respondents owned a car, and 42% had previously applied for a Section 8 housing voucher. The median adult age was 32 years.

5 Neighborhood Effects from MTO

We estimate LATE parameters using a Two-Stage Least Squares (TSLS) regression in which the first stage is:

$$D = Z\gamma + \eta, \tag{9}$$

and the second stage is:

$$Y = \widehat{D}\beta^{LATE} + \epsilon. \tag{10}$$

In these equations Y is the outcome of interest, D is a binary variable indicating receipt of a neighborhood treatment, and Z is experimental group status.

5.1 Defining Treatment

Many neighborhood characteristics are naturally defined as binary variables, and thus may be used as the treatment variable in the evaluation framework presented in Section 3. For example, one might be interested in whether residents believe a neighborhood is safe at night, or whether a child moves to a "good" school according to some criterion. But it is necessary to dichotomize a continuous treatment for it to fit into the evaluation framework in Section 3, which may cause Assumption 1 to fail to hold. Consider the example in which treatment is defined as moving to a neighborhood with a 20% poverty rate or less. A household that would move to a neighborhood with a poverty rate of 18% when not assigned treatment would be an always-taker under this definition of treatment. It is possible that such a household would be induced to move into a neighborhood with a lower poverty rate, say 10%, after being assigned treatment. However, given our binary definition of treatment, in this scenario the household would still be an always-taker. It might be the case that $Y(0)$ is not independent of Z if changes in treatment intensity across margins other than those defining the binary treatment affect outcomes.

The example of poverty rate shows a drawback of dichotomizing continuous treatments. On the other hand, defining treatment in such a way does help to focus attention on those margins of treatment believed to be important. We now consider three variables to help illustrate this point

We first investigate the consequences of dichotomizing a continuous measure of school quality. Consider the percentile ranking on test scores of schools in Chicago as illustrated in Table 5. With a first-stage F-statistic across all sites of 13, experimental group assignment is a strong instrument for school ranking if we use a continuous measure of school ranking. One line up in Table 5, however, shows that experimental group status becomes a weak instrument if we define school ranking as a binary variable indicating whether a child attended a school above or below the median school in their state.

Figure 14c shows the underlying distributions generating these statistics, and illustrates why dichotomizing a continuous treatment into a binary treatment can be a useful exercise. These CDFs show, first of all, that children in both the control and experimental groups attended the worst schools in Chicago. There is very little difference in the distributions of the control and experimental MTO youth up to their 60th percentiles. Between the 60th and 99th percentiles there is some difference in the distributions, but focusing on the vertical distance between the two CDFs, we see that most of this difference comes from moving children across margins near the 20th percentile school. It seems reasonable to assume the effects of moving MTO children to a school above the 20th percentile are negligible relative to moving children from this same population to a school above the median.

Another neighborhood characteristic measured continuously is the high school graduation rate of females within a census tract. Figure 4 shows the distribution of the US population in 2000 by the high school graduation rate of females over 25 in their census tract. The 10th percentile person lived in a neighborhood with a female high school graduation rate of 61.6 percent, and the graduation rates in the neighborhood of the 25th and 50th percentile individuals were, respectively, 73.2 and 83.2 percent.[13]

The national distribution in Figure 4 helps to put the MTO results in Figure 10 in context. Consider specifically Figure 10d, the distributions of the MTO treatment and experimental groups in Los Angeles. We can first see that the vast majority in the MTO control group lives in a neighborhood below the 10th percentile of the national distribution, which contrasts with Boston and Chicago. The next feature of Figure 10d we might notice is that MTO induced a large change in the experimental group distribution. We see a large area between the control and experimental CDFs, but most of this difference is between graduation rates of 35 and 75 percent. Moving people from a neighborhood with a graduation rate of 35 percent to a neighborhood with a 60 percent graduation rate moves them from a neighborhood *extremely* far in the left tail of the national distribution to a neighborhood that is still far in the left tail of the distribution. We see a much smaller vertical distance between the distributions when we look at the median graduation rate of 82.5. Thus Figure 10d indicates that MTO induced many families to move to neighborhoods with higher female high school graduation rates, but also that most of these changes took place somewhere in the long left tail of the national distribution.

The third continuous neighborhood characteristic we will consider is the share of households with own children under 18 that were single-headed. Figure 9 shows the MTO control and experimental distributions by site, and Figure 5 shows the national distribution. We can see in the figures that nearly all of the changes in the MTO experimental distribution took place to the right of the 75th percentile of the national distribution, 34.9%. In Baltimore, Chicago, and New York the majority of changes took place to the right of the 90th percentile of the national distribution, 48.6%, and in Boston and Los Angeles it was still the case that many of the changes were to the right of the 90th percentile. Similar to the two other continuous characteristics we have examined here, most of the changes in the share of single-headed households in a neighborhood that took place due to MTO are to be found far in the tail of the national distribution.

Looking at these distributions begs the natural question: What types of moves characterize a "strong" intervention? Reasonable people might disagree about what types of neighborhood characteristics are feasible targets for housing mobility programs. Such people might also disagree about what types of changes in neighborhood characteristics would be necessary to have large effects on individuals' outcomes. Together with the evaluation framework reviewed in this paper, the preceding discussion gives a constructive way to think about the consequences of adopting alternative definitions of a strong intervention.

Now consider how a dichotomized treatment fits into our evaluation framework by examining the experimentally induced variation in continuous neighborhood and school characteristics shown in Figures 7-17. These figures show the CDFs of continuous characteristics such as poverty rate, the rate of educational attainment, the labor force participation rate, and the student/teacher ratio. The vertical distance between the CDFs is the share of participants who were induced across

[13] The line in the figure is the graduation rate in the median census tract, not of the census tract in which the median resident lived. The median census tract line is shown in the figure because this is the cutoff used to define the binary treatments.

that margin due to the experiment. So for a binary treatment defined as inducing moves across a specific cutoff of each of these variables, the share of compliers is the vertical distance between the experimental and control CDFs. For example, Figure 7 shows that if we define treatment as moving to a neighborhood with a poverty rate of 20% or less, then 25.0% of residents in New York City were compliers. If we define treatment as moving to a neighborhood with a poverty rate of 10% or less, then 9.7% New York City residents were compliers.

When dichotomizing a continuous variable, the researcher is choosing a cutoff to classify those neighborhood characteristics qualifying as a neighborhood treatment. To come as close as possible to meeting Assumption 1, this cutoff should be chosen to define the margin across which moving will result in the largest effects. However, this goal must be balanced against choosing treatments the researcher might reasonably expect to observe. In order to balance these goals, the analysis in this paper defines high and low poverty and segregation in terms of MTO and Gautreaux program guidelines. The remaining neighborhood treatments are defined in terms of the median of all US census tracts. Using the median census tract to define neighborhood treatments is meant to represent moves to neighborhoods that are "good" along the dimension under consideration.

5.2 Treatments and Instrument Strength

Given binary neighborhood treatments defined in the previous section, we now investigate how strong of an instrument experimental group status in MTO was for inducing households into those treatments at the various MTO sites. This is a useful exercise because it recasts one controversy in the literature on MTO, which can be viewed as a debate regarding whether experimental group assignment was a weak instrument for various neighborhood treatments.[14]

We first gauge instrument strength visually by looking at Figures 7-13. One measure of the strength of the MTO intervention is the total area between the CDFs of the control and experimental groups. The measure we will use under our definition of neighborhood treatments is the share of compliers, which can be determined by looking at the vertical distance between the CDFs at the vertical lines in these figures, which represent either the program cutoff or median census tract characteristic in 2000 used to define treatment.

Using this visual criterion, we can first see that the largest changes in neighborhood characteristics due to the MTO intervention were related to the neighborhood poverty rate (Figure 7). The next largest changes were in the neighborhood female labor force participation rate (Figure 11). The other figures show remarkably little variation induced by MTO in many of the neighborhood characteristics believed to influence outcomes, including segregation, share of single-headed households, educational attainment, and the share of residents receiving public assistance income.

Figures 14-16 show there was remarkably little difference between the experimental and control

[14] Another approach would be to simply proceed in estimating LATE parameters using data from all of the sites, and then to make inference using tests whose properties have been established for arbitrarily weak instruments (Moreira (2009), Andrews et al. (2006)). But the analysis in this paper aims to paint a descriptive picture about the sign and magnitudes of various neighborhood effects rather than testing any specific hypothesis. As a result we are still concerned with the bias caused by weak instruments (Bound et al. (1995)).

groups in several measures of school quality. For example, essentially no children attended schools ranked above the median on statewide standardized exams, and thus only a very tiny minority of students could have been induced across this margin of school quality due to the experiment. Examining Figures 14 and 15, it appears that larger changes in class size and test score ranking were induced by Gautreaux than by MTO (Pages 131 and 162 of Rubinowitz and Rosenbaum (2000), respectively.).

Looking at the interventions by site rather than only by treatment, Figures 7-16 indicate that Boston, Los Angeles, and New York were the strongest interventions, and that Baltimore and Chicago were much weaker interventions.

We now follow a common practice in the literature for assessing the strength of an instrument and examine the F-statistic on the excluded instrument in the first-stage regression (Bound et al. (1995)). Tables 3 and 5 show the F-statistics of the excluded instrument (assigned treatment) in the first-stage regression (Equation 9) using several definitions of treatment (D). We will use these F-statistics to make formal statements about the experimentally-induced variation shown in Figures 7-16. We will call assigned treatment a weak instrument for treatment at a particular site if the F-statistic from the regression of treatment on assigned treatment is less than 10, a rule of thumb established in Staiger and Stock (1997) and used widely in the literature.[15]

First looking at adults, the results in Table 3 indicate that being assigned to the experimental group versus the control group is a very strong instrument for moving with an MTO voucher. However, these tables also show that the MTO intervention was, remarkably, a weak instrument at many sites for inducing changes in many of the other neighborhood characteristics believed to influence outcomes. At only one site each was experimental group status a strong instrument for moving to a neighborhood with a high rate of female high school graduates or BA holders, moving to a neighborhood in which a large share of employed females were employed in a high status occupation, moving to a neighborhood with low household rates of public assistance income, or moving to a neighborhood where police come when called. At no site was experimental group status a strong instrument for moving to an integrated neighborhood or moving to a neighborhood with a low rate of single-headed households. These data provide evidence against the idea that the neighborhood poverty rate is an adequate proxy for all of the neighborhood characteristics believed to influence outcomes.[16]

In addition to moving with an MTO voucher, experimental group status was also a strong instrument for moving to neighborhoods with low poverty, moving to neighborhoods with high female labor force participation rates, and moving to neighborhoods residents felt were safe at night. Experimental group status was a strong instrument for each of these treatments at either three or four sites, and these are the neighborhood treatments whose effects on adults will be estimated. Estimation will only use those sites at which experimental group status was judged to be a strong instrument for the neighborhood treatment under consideration.

[15] Although this need not be the only criterion used to determine the strength of an instrument (Cruz and Moreira (2005)), it is a rule of thumb widely used in the literature.

[16] In other words, this is evidence against the linear index assumption made in Kling et al. (2007a).

Looking at youth outcomes, Table 5 shows that being assigned to the experimental group is a weak instrument for understanding the effects of attending better schools in terms of standardized test scores, student/teacher ratios, time spent on homework, or several other measures of school quality. This is particularly striking because school quality can be considered an important measure of neighborhood quality for households with school-age children due to the way schools are financed in the US. We can also see in Table 5 that experimental group status was not a strong instrument for the collective efficacy in a youth's neighborhood, or for changing whether a youth had friends in their old neighborhood or whether members of their peer groups use drugs or carry weapons. Aside from poverty, neighborhood safety is the only other neighborhood or school characteristic for which experimental group status in MTO was a strong instrument at more than one site, an important point when considering LATE estimates on youth outcomes.

Although these results on instrument strength do not call into question any of the program effects obtained from the MTO data, Tables 3 and 5 do suggest that with the exception of poverty rate, the results from MTO are far from ideal evidence on causal effects of many of the neighborhood characteristics commonly suspected of influencing outcomes.

5.3 Assumption 1

Before proceeding to the estimation results, we first consider two possible violations of Assumption 1. The first violation comes from defining treatment in a binary way, and was discussed in depth in Section 5.1. Although the benefit of dichotomizing continuous treatments is forcing the researcher to define the margin at which they believe there will be the largest effects, it is not clear whether this benefit is outweighed by the potential violations of Assumption 1 it creates. This issue must be kept in mind when interpreting estimates. One way of addressing this issue is to define a variable intensity treatment rather than a binary treatment. Angrist and Imbens (1995) introduced a parameter that generalizes the LATE parameter to such a case where there is a discrete, multi-valued treatment, and Heckman et al. (2006) further develop this Average Causal Response (ACR) parameter within an ordered choice framework. However, while both of these frameworks could possibly be fruitful avenues for future research, both have strong limitations and will not be used here.[17]

The second possible violation of Assumption 1 comes from the fact that MTO changed a bundle of neighborhood characteristics. This violation can be seen by supposing that treatment D is defined in terms of the neighborhood poverty rate, but the outcome in question Y is also affected by neighborhood safety. If experimental group status Z induces changes in both neighborhood poverty rates and neighborhood safety, then it is likely that $Y(1)$ and $Y(0)$ will not be independent of Z.

Despite this violation of Assumption 1, we proceed with our identification strategy because it is difficult to imagine an instrumental variable identification scheme in which Assumption

[17] The key limitation of the Angrist and Imbens (1995) framework is the ability to interpret estimated parameters. Heckman et al. (2006) trade this limitation in favor of strong assumptions on the selection process.

1 is not violated. As one example, consider the well-known instrument originally proposed in Angrist and Krueger (1991), date of birth combined with compulsory schooling laws to instrument for educational attainment at a given age. Changing one's birth date changes the age at which one is eligible to drop out of school, and therefore induces variation in educational attainment at a given age. However, changing date of birth also changes the absolute age at which children enter school and their age relative to their classmates. It is difficult to say how large these effects are relative to the effects of attainment. But it is clear that this instrument changes multiple variables that are all likely to causally effect outcomes, creating a similar violation of Assumption 1.[18]

It is difficult to imagine an instrumental variable that induces changes in only one causal variable rather than a group of variables that causally effect the outcome of interest. This violation of Assumption 1 illustrates the limitations of focusing on exogenous variation in one causal variable when there are many variables causally effecting the outcome. Interpretation can be difficult when focusing on "the" effect of one causal variable while abstracting from the effects of other causal variables and their interactions.

5.4 Estimation Results

LATEs are estimated using data only from those sites at which assigned treatment is considered to be a strong instrument for the specific treatment under consideration, and only if experimental group status was a strong instrument at more than one site. Table 6 shows LATE estimates for adult outcomes. We see large and statistically significant effects on labor force participation rates for all definitions of treatment in the table. There are corresponding increases in both the employment and unemployment rates, but only the effects of poverty on unemployment are statistically significant.

A low neighborhood poverty rate, a high degree of neighborhood safety, and a high female labor force participation rate all have positive effects on income and negative effects on welfare benefits. Although these effects are of the expected sign, and many are economically significant, they are very imprecisely estimated. Only the effect of moving to a safe neighborhood on the household head's income is estimated precisely enough to be statistically significant.

The LATEs of neighborhood poverty rate on health are almost all statistically significant. And while only one of the LATEs of neighborhood safety and female labor force participation rate on health is statistically significant, the remaining effects are all of the expected sign. These effects are not too far from being statistically significant, and taken together they point to economically significant effects on health outcomes.

The share of compliers for the estimated LATEs is 21.8 percent of adults who moved from neighborhoods with poverty rates higher than 20 percent to neighborhoods with poverty rates below 20 percent, and an analogous 12.0 percent of adults who moved across the 10 percent margin. For neighborhood safety and female labor force participation the share of compliers is, respectively,

[18] This is ignoring other criticisms of this identification scheme caused by redshirting (Aliprantis (2011)), the non-random nature of birth date (Bound and Jaeger (2000)), and the weak correlation between date of birth and educational attainment (Bound et al. (1995)).

18.1 and 17.9 percent. Only neighborhood safety has a high share of always-takers; all of the other treatments have shares of always-takers under 15 percent, with the share for poverty rate less than 10 percent only 3.5 percent. Thus the share of compliers tends to be high relative to the share of always-takers, but also indicates the estimated LATEs are the average neighborhood effects on between 12.0 and 21.8 percent of the MTO sample.

The low share of always takers helps to illustrate how far in the tails of the national distributions were the neighborhood characteristics of the census tracts in which most MTO families were living. Of the entire nation's population, 81.5 and 52.9 percent lived in neighborhoods with less than 20 or 10 percent poverty rates, respectively, in 2000 (Figure 6). This compares with shares of always-takers at the sites used in estimation of only 14.5 and 3.5 percent, respectively.

Since MTO did not induce large changes in school quality along several measures, for youth outcomes we consider estimates of effects from neighborhood characteristics alone. Table 7 reports estimates of effects on youth outcomes related to problem behaviors, school outcomes, and health from moving to a low poverty neighborhood or a safe neighborhood. Nearly all of the estimates have very large standard errors and are thus difficult to interpret. One possible explanation is that these large standard errors are driven by heterogeneity in treatment effects. Since the literature shows strong heterogeneity in program effects on youth by gender, Table 8 shows a subset of LATE estimates for youth outcomes by gender. Although these estimates also have large standard errors, they are indicative of heterogeneity in youth LATEs by gender.

Very few of the effects on youth outcomes in Tables 7 and 8 are statistically different from zero, and there are two explanations for this. The first is that changes to neighborhood characteristics alone are not enough to effect outcomes if they are not also combined with improvements to school quality. This explanation is consistent with previous findings in the literature (Oreopoulos (2003)), and need not preclude the possibility that schools alone or schools together with neighborhoods or social programs (Dobbie and Fryer (2011)) can have large effects on youth outcomes. A second explanation is that changes in neighborhood characteristics effect girls and boys differently. Both of these explanations merit further attention, but the lack of improvement in school quality must be the leading explanation for the absence of strong effects on youth outcomes from MTO.

In addition to the results on instrument strength presented in Section 5.2, the difference in magnitudes between LATE and ITT and TOT estimates could help to resolve some of the controversy in the literature about the appropriate interpretation of the results from MTO. For example, the LATEs on employment range from 10 to 29 percentage points in Table 6, which contrast strongly with the ITT and TOT estimates of 1.5 and 3.3 percentage points, respectively, reported in Table F3 of Kling et al. (2007b). Similarly, the LATEs of moving to a safe neighborhood are \$4,743 for earnings in 2001 and \$6,156 for household head's income, compared with ITT and TOT estimates on earnings in 2001, respectively, of –\$287 and –\$612 as reported in Table F4 of Kling et al. (2007b). These results suggest neighborhood effects are of a larger magnitude than the program effects from MTO.

6 Discussion

6.1 Site Heterogeneity and the Design of Housing Mobility Programs

Complementing the data considered in Section 5.2, Tables 10 and 11 show large differences in control means and program effects across sites.[19] Since investigation of this heterogeneity in program effects from MTO could be a fruitful avenue for improving both our understanding of neighborhood effects and the design of future housing mobility programs, some speculative hypotheses for explaining differences across sites are considered here.

We can see from Table 10 that positive adult labor market outcomes are driven by the effects in New York. The positive program effects on labor market outcomes in New York could be driven by the fact that income was lowest there, as were employment and labor force participation rates. These effects could also be driven by the fact New York and Los Angeles had the least safe neighborhoods as judged by participants, and these sites also had the largest increases in safety due to the program. Transportation could also help to explain differences across sites. Car ownership was an important issue for Gautreaux movers (Polikoff (2006), p 222), and New York's public transportation could give residents an advantage in accessing local labor markets. Another feature of New York's implementation of MTO is that it was the one site at which none of the participants lived in public housing that was a part of HOPE VI.

A final suspicion is that the positive labor market results in New York were driven by the high share of hispanic participants at that site, but this was not the case. Estimated ITT effects (standard errors) on household head's total income, employment, and labor force participation rate for blacks in New York City are 3,799 (1,332), 0.11 (0.06), and 0.16 (0.06). These ITT effects for non-black participants in New York City are 1,914 (1,087), 0.10 (0.06), and 0.05 (0.06).

The program effects in Baltimore compare unfavorably with those from the other sites, and racial segregation is one possible explanation for these outcomes. Consider first Figure 18 and Table 9, which show decennial census data from the NHGIS indicating that African Americans living in MTO cities in 1990 were dramatically more segregated from whites than other minority groups. Table 9 and Figures 19a and 19c show that this difference was most pronounced in Baltimore and Chicago, the sites at which MTO participants were almost entirely African American (Table 10), and also the sites with the weakest interventions as discussed in Section 5.2. The median black person in Baltimore and Chicago lived in a neighborhood almost completely devoid of any whites in 1990. Figure 19 illustrates there are large shares of African Americans in the other MTO sites who also lived in neighborhoods with extremely few whites even in 1990. Figure 8 confirms that MTO participants in both the experimental and control groups were those living in such highly segregated neighborhoods.

Turning our attention specifically to Baltimore, we see that it was by far the weakest of the MTO interventions. Baltimore was the only site at which experimental group status was not a strong instrument for moving to a low poverty neighborhood under any definition (Table 3), and it induced

[19] Note that Tables 6–8 can be anticipated by combining Table 10 with Tables 3 and 5.

the least moves to safe neighborhoods (Table 10). Baltimore had negative ITT effects on income and labor force participation rates, and ITT effects of increased arrest rates for both males *and* females (Table 10). These facts are especially noteworthy because Baltimore was also the site at which there was a strong, hostile response to the program along racial lines (Rubinowitz and Rosenbaum (2000), p 184.). Racial segregation has been found to be the most important factor explaining arrest rates of MTO youth for violent crimes (Ludwig and Kling (2007)), and this evidence indicates racial segregation could also be an important factor in explaining other outcomes of interest to researchers. In terms of designing housing mobility policies, the experience in Baltimore refocuses attention on racial segregation. This experience also underscores the importance of effectively communicating the size and concentration of movers in receiving communities, as well as more research being conducted related to the effects of desegregation policies on receiving communities, similar to that already conducted on effects from HOPE VI (Hartley (2010)) or Boston's Metco program (Angrist and Lang (2004)).

Improving school quality is an obvious goal of housing mobility programs, but there were not large improvements to school quality made through MTO. Evidence from HOPE VI (Jacob (2004)) and school choice programs (Cullen et al. (2006), Hastings and Weinstein (2008)) indicates that information may be an important part of this process. As discussed in the literature, school choice may complicate the design of housing mobility programs. Thirty percent of MTO control group children in Chicago and Los Angeles were attending magnet schools (Sanbonmatsu, p 684). Since schools are likely to play a large role in improving youth outcomes, one possibility for the design of effective housing voucher programs would be to provide a voucher conditional on children attending a school meeting some criterion, regardless of whether the parents use the voucher to move. Such an approach to the design of housing voucher programs might look more like a Conditional Cash Transfer program like Progresa (Todd and Wolpin (2006)) than a traditional housing mobility program. One program following this model is the St. Paul Early Childhood Scholarship Program, which gives scholarships to help families access early care and education programs, and requires that families use these scholarships at programs meeting certain quality rating standards.

7 Conclusion

This paper presented a new perspective on the interpretation of results from MTO, especially as they relate to neighborhood effects. The paper provided a review of recent advances in the program evaluation literature in order to make a clear distinction between the interpretation of Intent to Treat (ITT) and Treatment on the Treated (TOT) parameters as program effects and Local Average Treatment Effect (LATE) parameters as neighborhood effects. Due to the nature of the LATE, this distinction helped to clarify that results from MTO are only informative about a small subset of neighborhood effects of interest. The evaluation framework was also used to emphasize that since the LATE is defined by the subgroup of compliers, different instruments will result in different LATE parameters if they induce different subpopulations to select into treatment

(Heckman (1997)). An important implication was that alternative housing mobility programs designed to induce moves to neighborhoods with characteristics in addition to or in lieu of low poverty might induce larger effects than MTO.

After this review of the literature, the paper used experimental group status in MTO as an instrumental variable to estimate the LATEs of various neighborhood characteristics. A first step in this process was to investigate at which sites experimental group status was a strong instrument for neighborhood treatments. Tests for instrument strength showed that MTO induced large changes in neighborhood poverty rates. However, it was also shown that MTO induced remarkably little variation in many of the other neighborhood and school characteristics believed to influence outcomes and that much of this variation was confined to the tails of these characteristics' national distributions. Such characteristics include school quality, as well as the female high school graduation rate and the share of single-headed households in participants' neighborhood of residence. This investigation also showed substantial differences in outcomes across MTO sites, pointing to the investigation of heterogeneity in program effects from MTO as a fruitful direction for future research.

LATE estimates at sites where experimental group status was a strong instrument were consistent with prevailing theories of neighborhood effects. Moves to neighborhoods with low poverty rates, a high degree of personal safety, or a high female labor force participation rate were all associated with increases in labor force participation. Moves to low poverty and safe neighborhoods were also associated with improved health outcomes. And although improvements in labor market outcomes such as employment and income coming from moves to such neighborhoods were not estimated precisely enough to be statistically significant, these effects were improvements and they were of large magnitudes.

References

Aliprantis, D. (2011). Redshirting, compulsory schooling laws, and educational attainment. *Journal of Educational and Behavioral Statistics* (doi: 10.3102/1076998610396885). Forthcoming.

Aliprantis, D. and M. Zenker (2011). Concentrated poverty. *Mimeo., Federal Reserve Bank of Cleveland*.

Andrews, D. W. K., M. J. Moreira, and J. H. Stock (2006). Optimal two-sided invariant similar tests for instrumental variables regression. *Econometrica 74*(3), 715–752.

Angrist, J. D. and G. W. Imbens (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association 90*(430), 431–442.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996). Identification of causal effects using Instrumental Variables. *Journal of the American Statistical Association 91*(434), 444–455.

Angrist, J. D. and A. B. Krueger (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics 106*(4), 979–1014.

Angrist, J. D. and K. Lang (2004). Does school integration generate peer effects? Evidence from Boston's Metco program. *The American Economic Review 94*(5), 1613–1634.

Bound, J. and D. A. Jaeger (2000). Do compulsory school attendance laws alone explain the association between quarter of birth and earnings? In S. W. Polachek (Ed.), *Worker Well Being*, Volume 19, pp. 83–108. Research in Labor Economics.

Bound, J., D. A. Jaeger, and R. M. Baker (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogeneous explanatory variable is weak. *Journal of the American Statistical Association 90*(430), 443–450.

Burdick-Will, J., J. Ludwig, S. W. Raudenbush, R. J. Sampson, L. Sanbonmatsu, and P. Sharkey (2010). Converging evidence for neighborhood effects on children's test scores: An experimental, quasi-experimental, and observational comparison. *Mimeo., Brookings Institution*.

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.

Clark, W. A. V. (2008). Reexamining the Moving to Opportunity study and its contribution to changing the distribution of poverty and ethinic concentration. *Demography 45*(3), 515–535.

Cruz, L. M. and M. J. Moreira (2005). On the validity of econometric techniques with weak instruments: Inference on returns to education using compulsory school attendance laws. *The Journal of Human Resources 40*(2), 393–410.

Cullen, J. B., B. A. Jacob, and S. Levitt (2006). The effect of school choice on participants: Evidence from randomized lotteries. *Econometrica 74*(5), pp. 1191–1230.

DeLuca, S. and J. E. Rosenbaum (2003). If low-income blacks are given a chance to live in white neighborhoods, will they stay? Examining mobility patterns in a quasi-experimental program with administrative data. *Housing Policy Debate 14*(3), 305–345.

Dobbie, W. and R. G. Fryer, Jr. (2011). Are high-quality schools enough to increase achievement among the poor? Evidence from the Harlem Children's Zone. *American Economic Journal: Applied Economics 3*(3), 158–187.

Durlauf, S. N. (2004). Neighborhood Effects. In J. V. Henderson and J. E. Thisse (Eds.), *Handbook of Regional and Urban Economics*, Volume 4. Elsevier.

Garces, E., D. Thomas, and J. Currie (2002). Longer-term effects of Head Start. *The American Economic Review 92*(4), 999–1012.

Goering, J. (2003). The impacts of new neighborhoods on poor families: Evaluating the policy implications of the Moving to Opportunity demonstration. *Economic Policy Review 9*(2).

Hartley, D. A. (2010). Blowing it up and knocking it down: The effect of demolishing high concentration public housing on crime. *Federal Reserve Bank of Cleveland Working Paper 10-22*.

Hastings, J. S. and J. M. Weinstein (2008). Information, school choice, and academic achievement: Evidence from two experiments. *Quarterly Journal of Economics 123*(4), 1373–1414.

Heckman, J. J. (1996). Randomization as an Instrumental Variable. *The Review of Economics and Statistics 78*(2), pp. 336–341.

Heckman, J. J. (1997). Instrumental Variables: A study of implicit behavioral assumptions used in making program evaluations. *Journal of Human Resources 32*(3), 441–462.

Heckman, J. J. (2010). Building bridges between structural and program evaluation approaches to evaluating policy. *Journal of the Economic Literature 48*(2), 356–398.

Heckman, J. J. and P. A. LaFontaine (2010). The American high school graduation rate: Trends and levels. *The Review of Economics and Statistics 92*(2), 244–262.

Heckman, J. J. and S. Urzúa (2010). Comparing IV with structural models: What simple IV can and cannot identify. *Journal of Econometrics 156*(1), 27–37.

Heckman, J. J., S. Urzúa, and E. Vytlacil (2006). Understanding Instrumental Variables in models with essential heterogeneity. *The Review of Economics and Statistics 88*(3), 389–432.

Heckman, J. J. and E. Vytlacil (2005). Structural equations, treatment effects, and econometric policy evaluation. *Econometrica 73*(3), 669–738.

Heckman, J. J. and E. J. Vytlacil (2000). The relationship between treatment parameters within a latent variable framework. *Economics Letters 66*, 33–39.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association 81*(396), 945–960.

Imbens, G. W. and J. D. Angrist (1994). Identification and estimation of Local Average Treatment Effects. *Econometrica 62*(2), 467–475.

Jacob, B. A. (2004). Public housing, housing vouchers, and student achievement: Evidence from public housing demolitions in Chicago. *The American Economic Review 94*(1), 233–258.

Jargowsky, P. A. (1997). *Poverty and Place: Ghettos, Barrios, and the American City.* New York: Russell Sage Foundation.

Jargowsky, P. A. (2003). *Stunning Progress, Hidden Problems: The Dramatic Decline of Concentrated Poverty in the 1990s.* Brookings Institute.

Keane, M. P. and K. I. Wolpin (2000). Eliminating race differences in school attainment and labor market success. *Journal of Labor Economics 18*(4), 614–652.

Keels, M., G. J. Duncan, S. Deluca, R. Mendenhall, and J. Rosenbaum (2005). Fifteen years later: Can residential mobility programs provide a long-term escape from neighborhood segregation, crime, and poverty? *Demography 42*(1), pp. 51–73.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007a). Experimental analysis of neighborhood effects. *Econometrica 75*(1), 83–119.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007b). Supplement to "Experimental analysis of neighborhood effects": Web appendix. *Econometrica 75*(1), 83–119.

Kling, J. R., J. Ludwig, and L. F. Katz (2005). Neighborhood effects on crime for female and male youths: Evidence from a randomized housing voucher experiment. *The Quarterly Journal of Economics 120*(1), 87–130.

LaLonde, R. (1995). The promise of public sector sponsored training programs. *Journal of Economic Perspectives 9*(2), 149–168.

Ludwig, J. (2010). Improving the life chances of disadvantaged children. *NBER Reporter* (3), 6–8.

Ludwig, J. and J. R. Kling (2007). Is crime contagious? *Journal of Law and Economics 50*(3), 491–518.

Ludwig, J., H. F. Ladd, G. J. Duncan, J. Kling, and K. M. O'Regan (2001). Urban poverty and educational outcomes [with comments]. *Brookings-Wharton Papers on Urban Affairs*, pp. 147–201.

Ludwig, J., J. B. Liebman, J. R. Kling, G. J. Duncan, L. F. Katz, R. C. Kessler, and L. Sanbonmatsu (2008). What can we learn about neighborhood effects from the Moving to Opportunity experiment? *American Journal of Sociology 114*(1), 144–188.

Ludwig, J. and D. L. Miller (2007). Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *The Quarterly Journal of Economics 122*(1), 159–208.

Mendenhall, R., S. DeLuca, and G. Duncan (2006). Neighborhood resources, racial segregation, and economic mobility: Results from the Gautreaux program. *Social Science Research 35*(4), 892–923.

Minnesotra Population Center (2004). *National Historical Geographic Information System* (Prerelease Version 0.1 ed.). Minneapolis, MN: University of Minnesota. http://www.nhgis.org.

Moreira, M. J. (2009). Tests with correct size when instruments can be arbitrarily weak. *Journal of Econometrics 152*(2), 131–140.

NCHS (2009). Health, United States, 2008. National Center for Health Statistics. Hyattsville, MD.

Neal, D. A. and W. R. Johnson (1996). The role of premarket factors in black-white wage differences. *Journal of Political Economy 104*(5), 869–895.

Oreopoulos, P. (2003). The long-run consequences of living in a poor neighborhood. *The Quarterly Journal of Economics 118*(4), pp. 1533–1575.

Orr, L. L., J. D. Feins, R. Jacob, E. Beecroft, L. Sanbonmatsu, L. F. Katz, J. B. Liebman, and J. R. Kling (2003). *Moving to Opportunity: Interim Impacts Evaluation.* Washington, DC: US Department of Housing and Urban Development, Office of Policy Development and Research.

Pettit, B. and B. Western (2004). Mass imprisonment and the life course: Race and class inequality in US incarceration. *American Sociological Review 69*, 151–169.

Polikoff, A. (2006). *Waiting for Gautreaux.* Northwestern University Press.

Reardon, S. (2008). Thirteen ways of looking at the black-white test score gap. Stanford Institute for Research on Education Policy & Practice, Working Paper 2008-08.

Rosenbaum, J. E. (1995). Changing the geography of opportunity by expanding residential choice: Lessons from the Gautreaux program. *Housing Policy Debate 6*(1), 231–269.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology 66*(5), 688–701.

Rubinowitz, L. S. and J. E. Rosenbaum (2000). *Crossing the Class and Color Lines: From Public Housing to White Suburbia.* University of Chicago Press.

Sampson, R. J. (2008). Moving to inequality: Neighborhood effects and experiments meet social structure. *American Journal of Sociology 114*(1), 189–231.

Sanbonmatsu, L., J. R. Kling, G. J. Duncan, and J. Brooks-Gunn (2006). Neighborhoods and academic achievement: Results from the Moving to Opportunity experiment. *The Journal of Human Resources 41*(4), 649–691.

Staiger, D. and J. H. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica 65*(3), pp. 557–586.

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. *The American Economic Review 96*(5), 1384–1417.

Votruba, M. E. and J. R. Kling (2009). Effects of neighborhood characteristics on the mortality of black male youth: Evidence from Gautreaux, Chicago. *Social Science & Medicine 68*(5), 814–823.

Wilson, W. J. (1987). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy.* University of Chicago.

Wilson, W. J. (1996). *When Work Disappears: The World of the New Urban Poor.* Knopf.

Figures

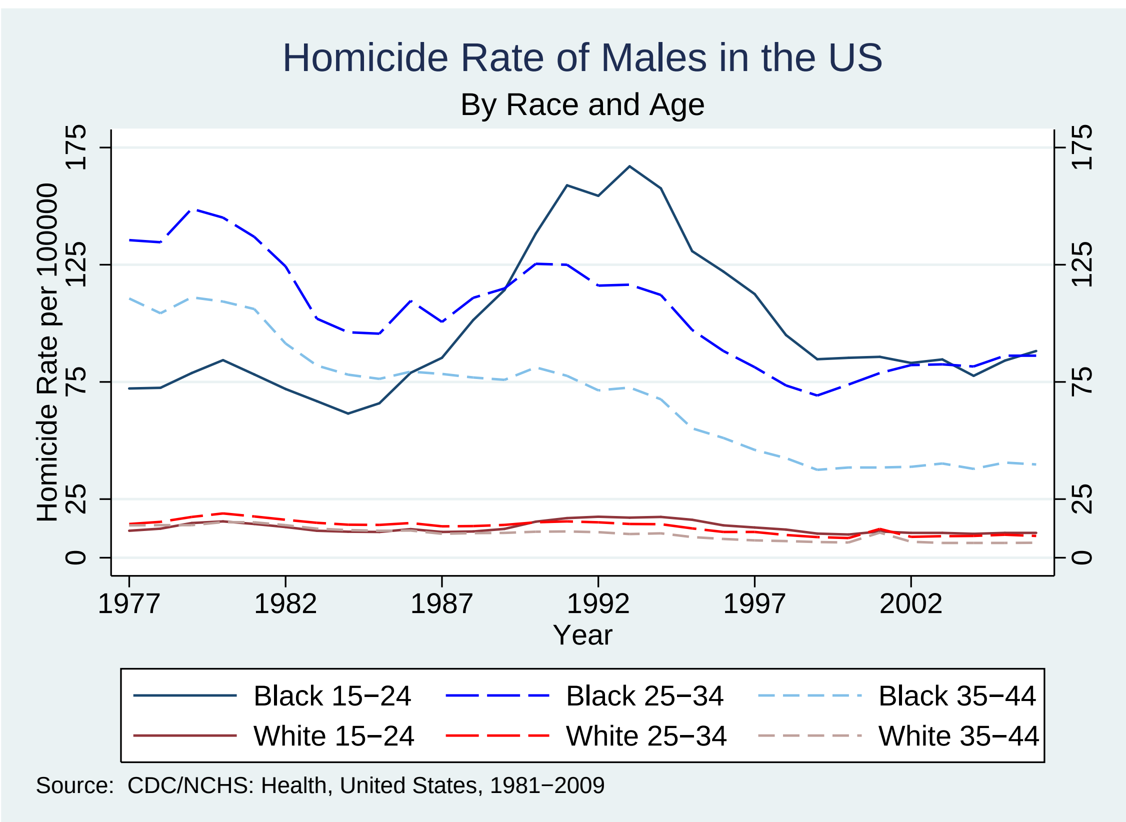


Figure 1: Homicide Rate of Males

Figure 2a: Type 1 Households: $\mu_1^*(Z=0) < \mu_1^*(Z=1)$

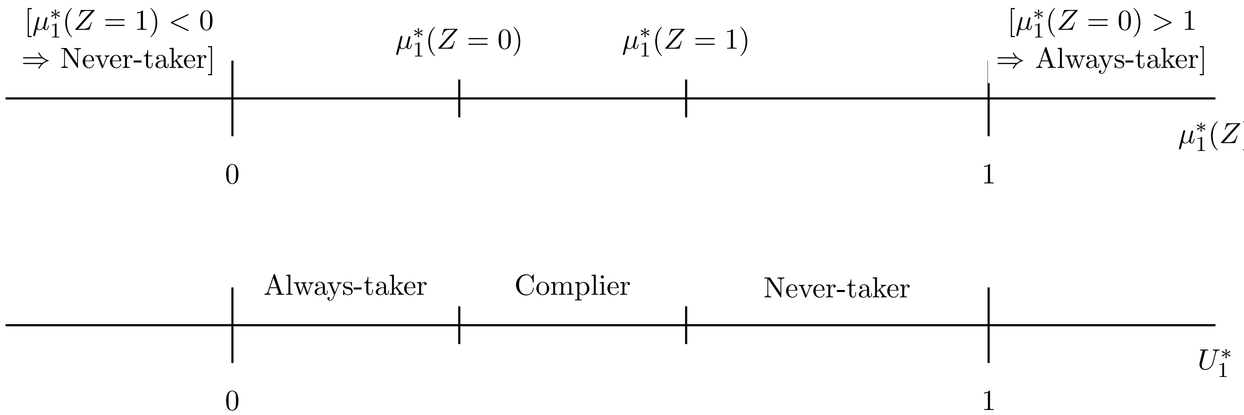
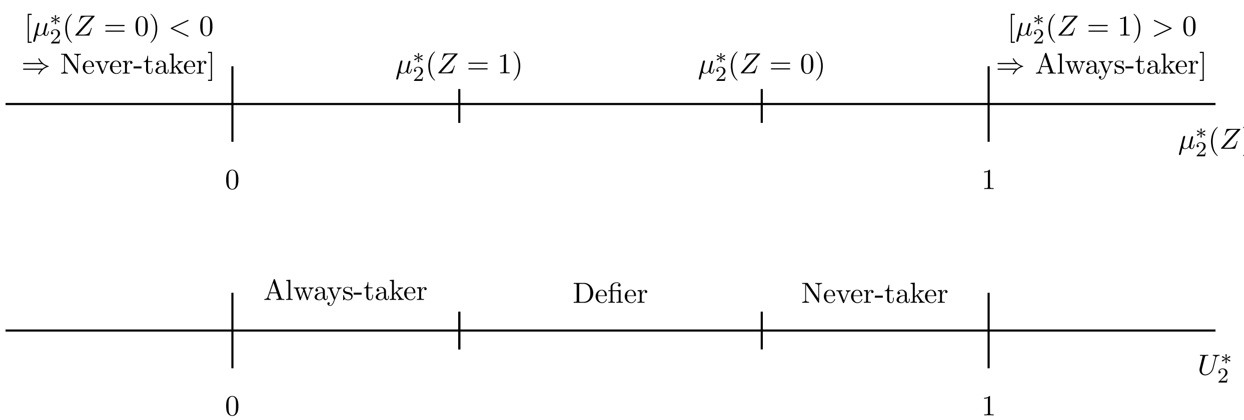


Figure 2b: Type 2 Households: $\mu_2^*(Z=1) < \mu_2^*(Z=0)$



$$D^* = \mu^*(Z) - U^*, \qquad D = 1\{D^* \geq 0\}, \qquad U^* \sim U[0,1].$$

Figure 2: Selection into Treatment

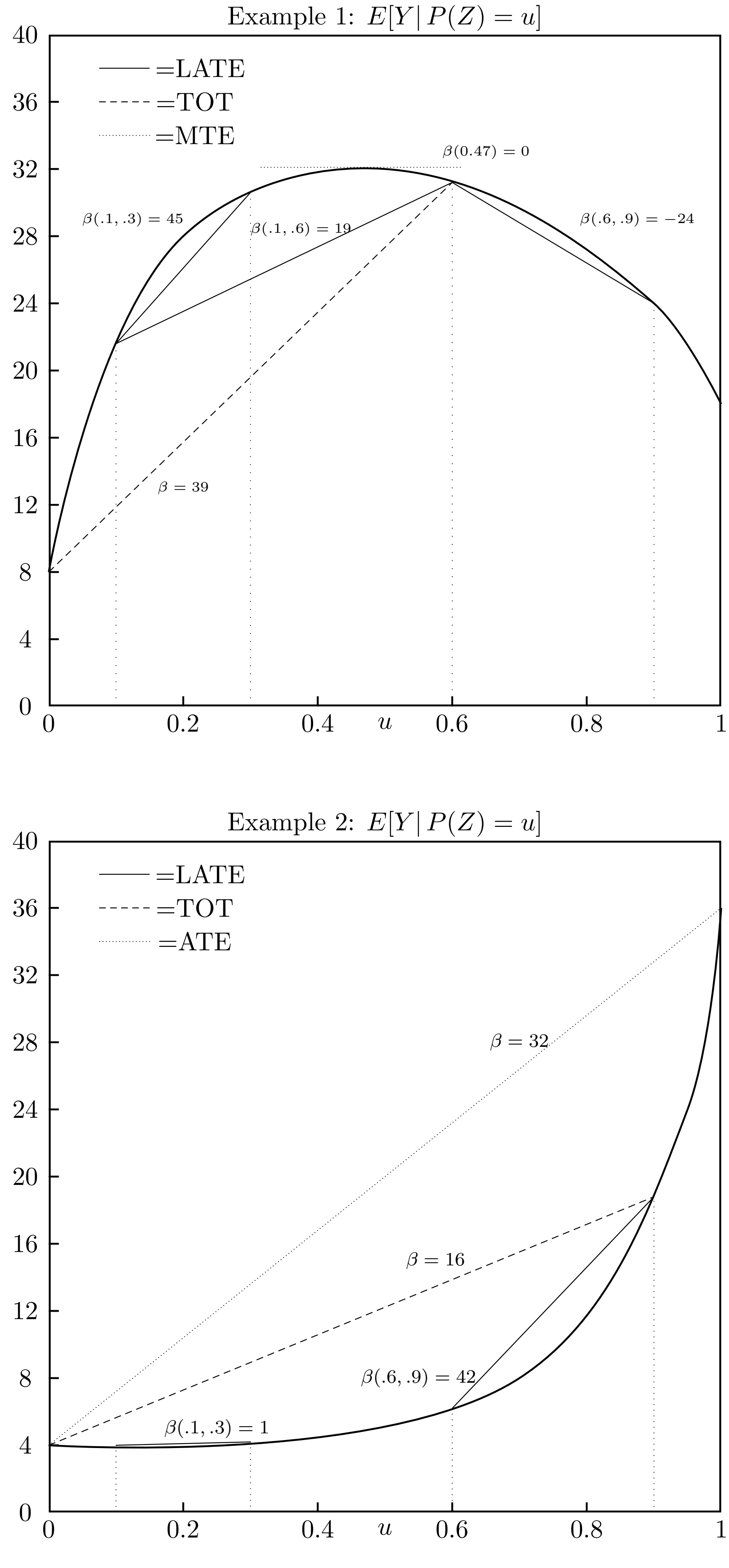


Figure 3: Some Example Treatment Effects

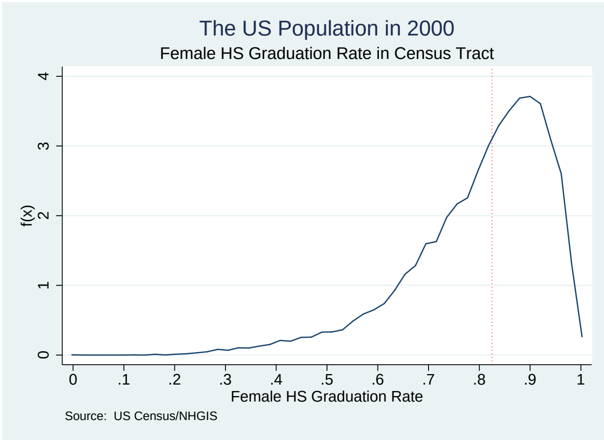


Figure 4: The US Population in 2000

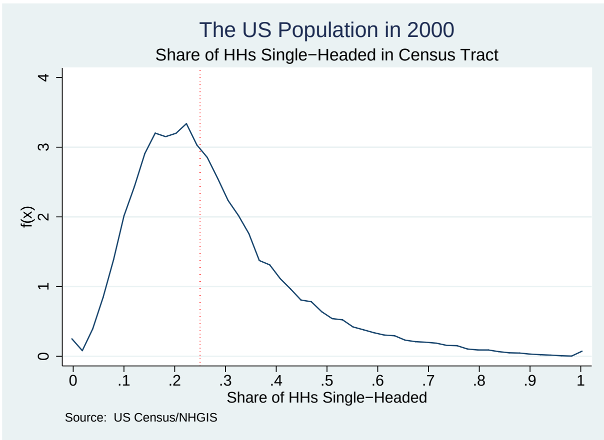


Figure 5: The US Population in 2000

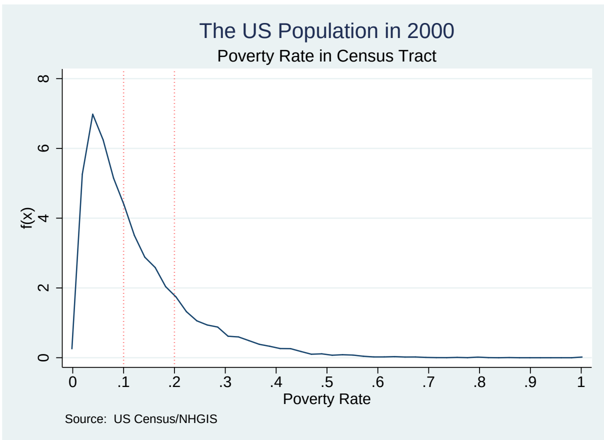
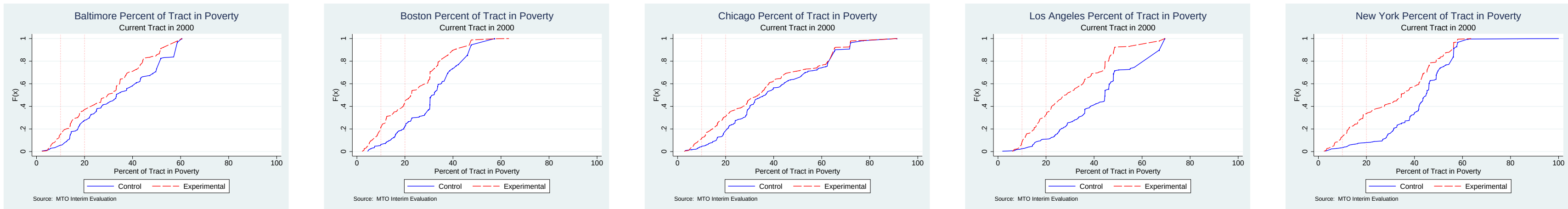
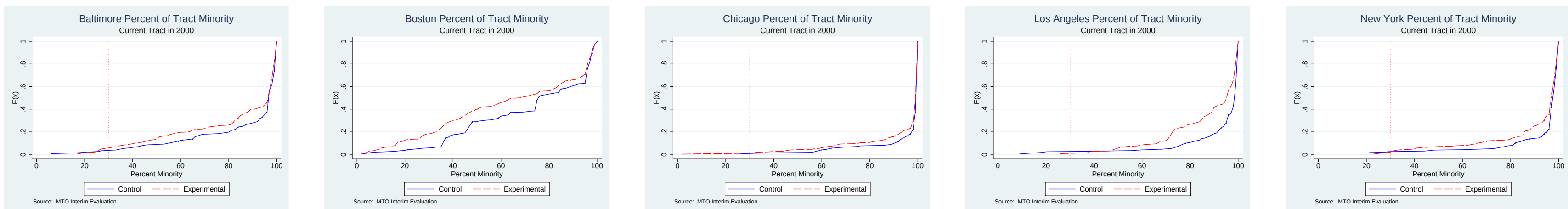


Figure 6: The US Population in 2000



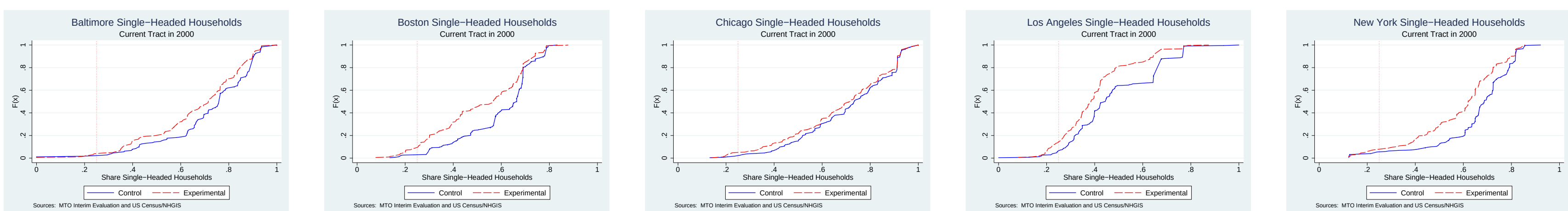
(a) Baltimore (b) Boston (c) Chicago (d) Los Angeles (e) New York City

Figure 7: Poverty Rate in 2002 Tract of Residence (Measured in 2000, by Site)



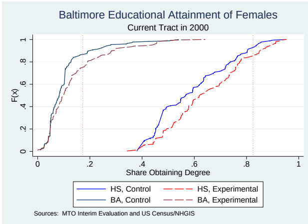
(a) Baltimore (b) Boston (c) Chicago (d) Los Angeles (e) New York City

Figure 8: Percent Minority in 2002 Tract of Residence (Measured in 2000, by Site)

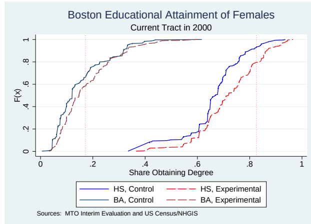


(a) Baltimore (b) Boston (c) Chicago (d) Los Angeles (e) New York City

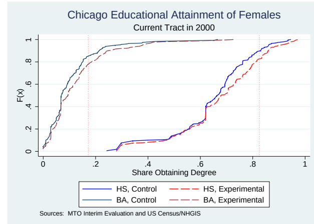
Figure 9: Share Single-Headed Households in 2002 Tract of Residence (Measured in 2000, by Site)



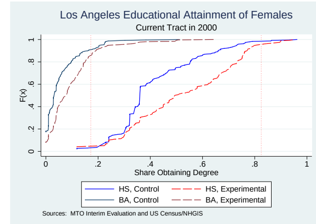
(a) Baltimore



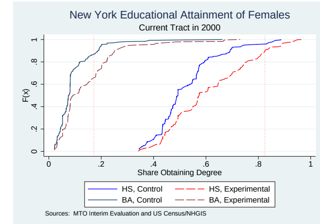
(b) Boston



(c) Chicago

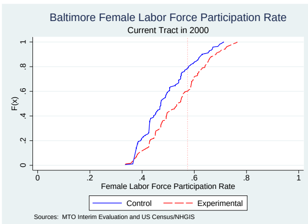


(d) Los Angeles

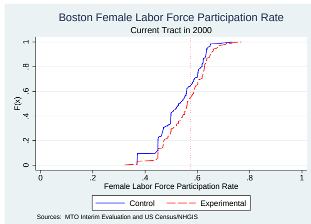


(e) New York City

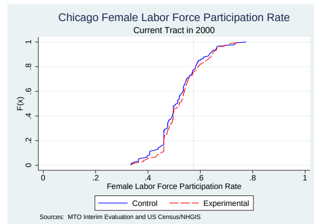
Figure 10: Female Educational Attainment in 2002 Tract of Residence (Measured in 2000, by Site)



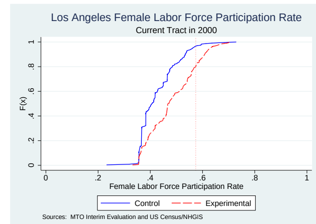
(a) Baltimore



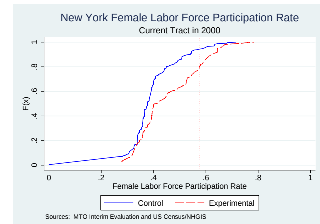
(b) Boston



(c) Chicago

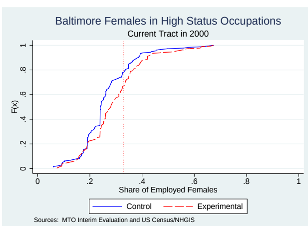


(d) Los Angeles

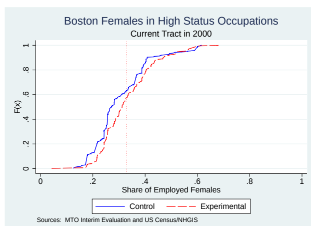


(e) New York City

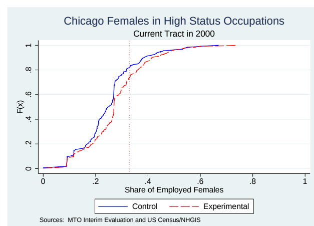
Figure 11: Female Labor Force Participation Rate in 2002 Tract of Residence (Measured in 2000, by Site)



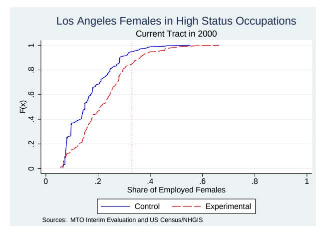
(a) Baltimore



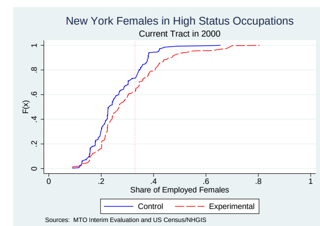
(b) Boston



(c) Chicago

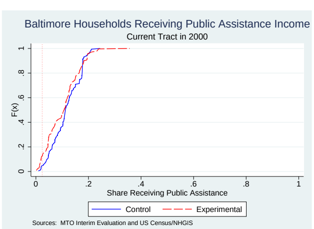


(d) Los Angeles

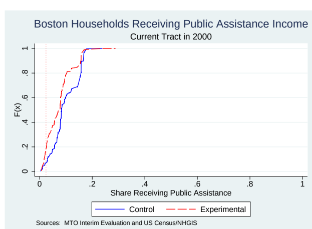


(e) New York City

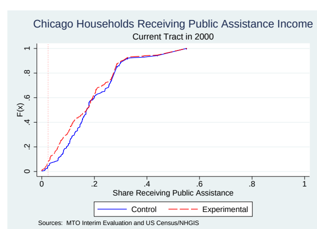
Figure 12: Share of Employed Females in High Status Occupations in 2002 Tract of Residence (Measured in 2000, by Site)



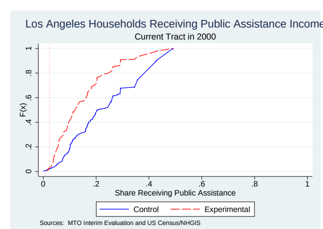
(a) Baltimore



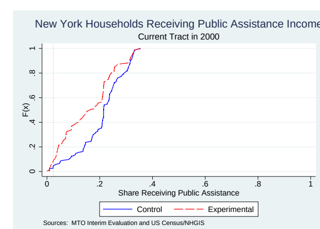
(b) Boston



(c) Chicago

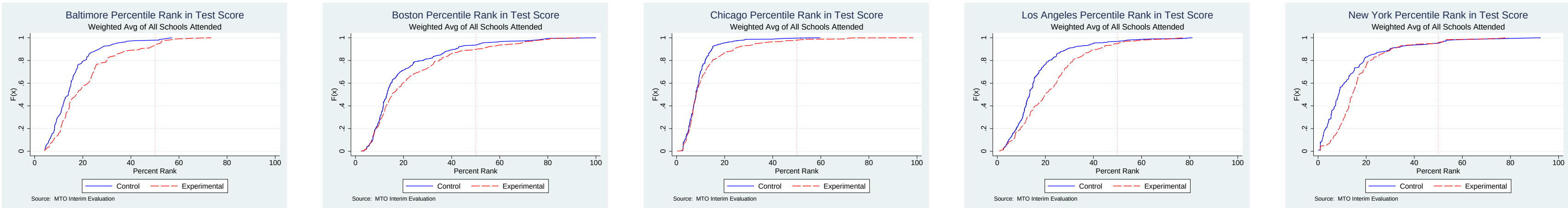


(d) Los Angeles



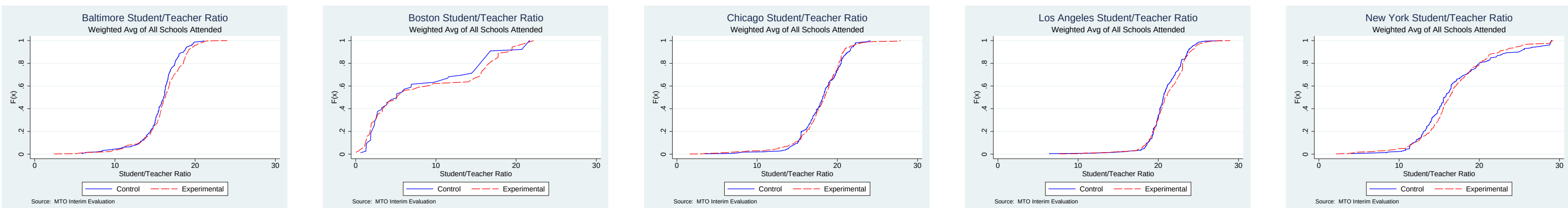
(e) New York City

Figure 13: Share of Residents Receiving Public Assistance Income in 2002 Tract of Residence (Measured in 2000, by Site)



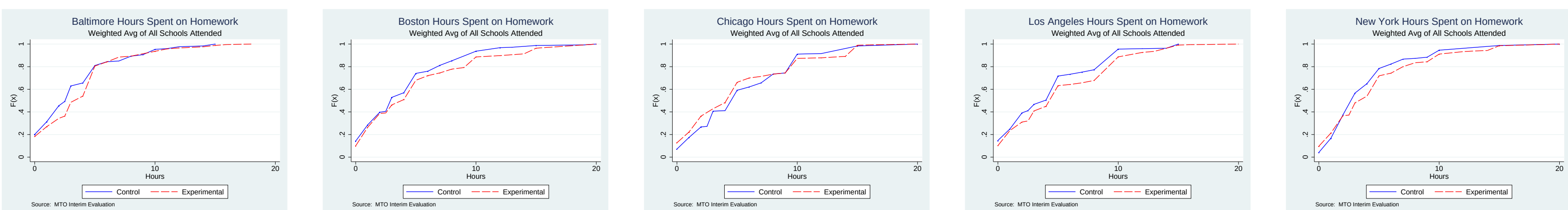
(a) Baltimore (b) Boston (c) Chicago (d) Los Angeles (e) New York City

Figure 14: School Ranking on State Tests, Weighted Average Percentile over all Schools Attended (by Site)



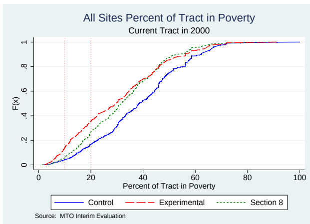
(a) Baltimore (b) Boston (c) Chicago (d) Los Angeles (e) New York City

Figure 15: Student/Teacher Ratio, Weighted Average over all Schools Attended (by Site)

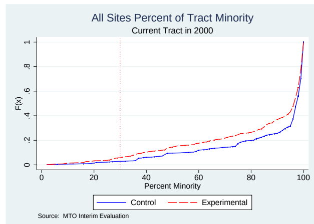


(a) Baltimore (b) Boston (c) Chicago (d) Los Angeles (e) New York City

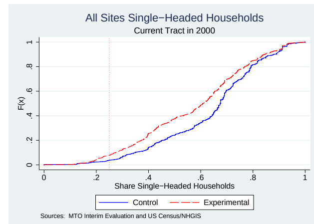
Figure 16: Hours/Week Spent on Homework, Weighted Average over all Schools Attended (by Site)



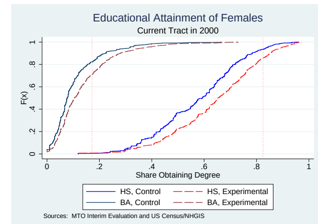
(a) Poverty Rate



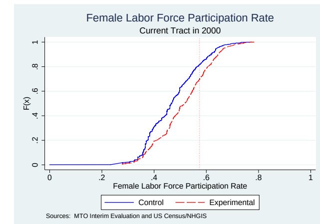
(b) Share Minority



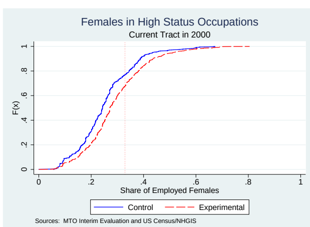
(c) Share Single-Headed HHs



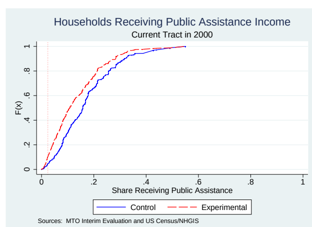
(d) Female Ed Attainment



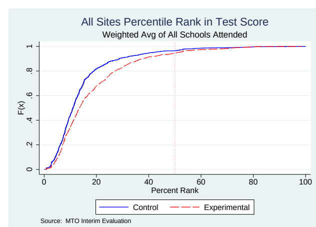
(e) Female LFP



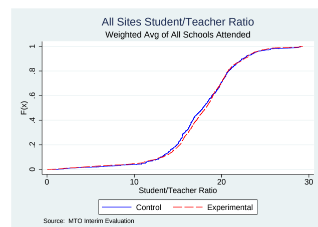
(f) High Status Occupations



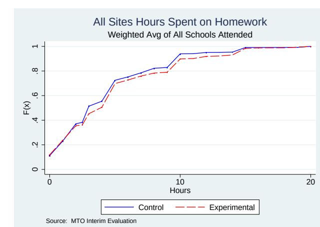
(g) HHs w Pub Assist Income



(h) School Ranking



(i) Student/Teacher Ratio



(j) Time Spent on Hmwk

Figure 17: Changes in Neighborhood and School Characteristics (All Sites Together)

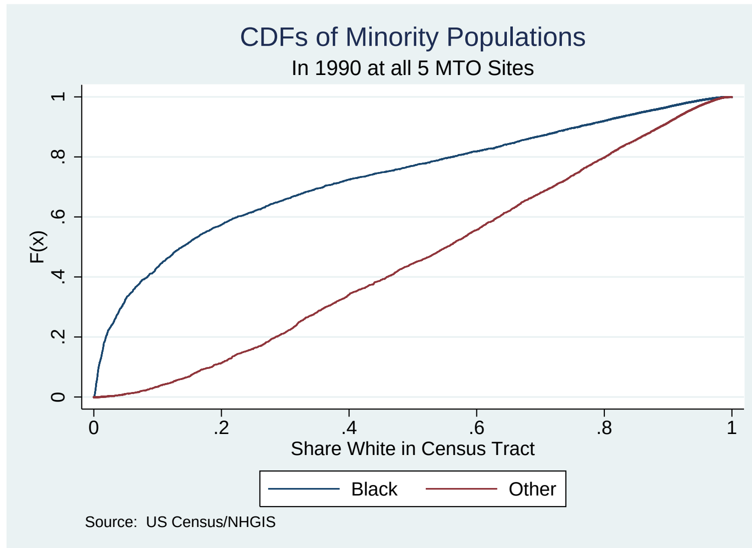
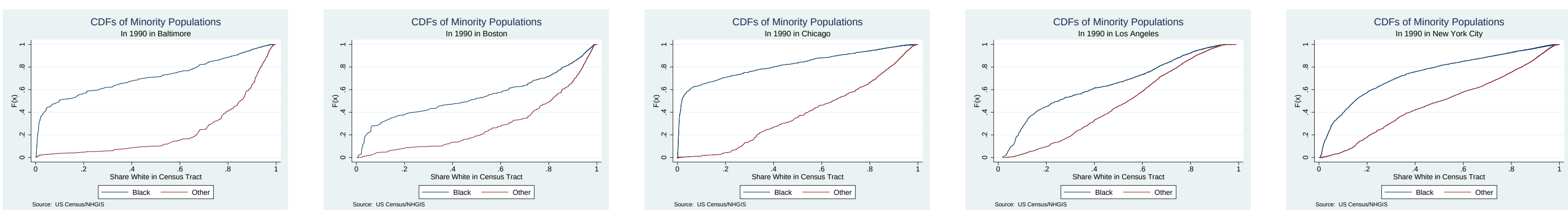


Figure 18: CDFs of Black and Other Minority Populations in 1990 (All MTO Cities Combined)



(a) Baltimore (b) Boston (c) Chicago (d) Los Angeles (e) New York City

Figure 19: CDFs of Black and Other Minority Populations in 1990 (by Site)

Tables

Table 1: D(Z): Treatment as a Function of Assigned Treatment

D(Z)		D(0)	
	D	0	1
D(1)	0	Never-taker	Defier
	1	Complier	Always-taker

Table 2: Descriptive Statistics of Sample Adults at Baseline (Percent)

Variable	Mean	SE
Receiving AFDC/TANF	75.08	0.67
HS Diploma	37.73	0.75
GED	18.55	0.60
Working for Pay	25.49	0.67
Nbd Streets Not Safe at Night	49.76	0.77
Applied for Section 8 Before	42.09	0.76
African American	62.75	0.75
Hispanic	30.17	0.71
White or Other	7.08	0.40
HH Member Owns a Car	15.98	0.57
HH Member Disabled	16.05	0.57
HH Member Victim of Crime	41.89	0.76

Table 3: F-Statistics: Adults
2000 Census Tract Characteristics, Tract of Residence in 2001

Binary Treatment	Baltimore	Boston	Chicago	LA	NYC	All Sites
Program						
Moving w/ MTO voucher	**178.7**	**170.7**	**173.2**	**294.7**	**183.0**	**909.4**
Neighborhood Characteristics						
Percent in poverty $\leq$ 20%	2.34	**24.28**	**11.67**	**27.16**	**36.68**	**89.09**
Percent in poverty $\leq$ 10%	6.64	**36.31**	5.21	**54.24**	**34.97**	**101.13**
Percent minority $\leq$ 30%	0.67	9.89	0.55	0.93	0.04	5.55
Female HS Grad Rate $\geq$ 83%	3.45	**15.26**	5.09	3.84	5.18	**30.84**
Female BA Attain Rate $\geq$ 17%	6.62	2.32	5.13	2.89	**23.08**	**32.93**
Female LFP $\geq$ 57%	**10.38**	3.16	0.26	**28.85**	**20.81**	**34.33**
Female High Stat Occ $\geq$ 33%	4.55	1.27	6.02	**12.91**	4.30	**19.53**
HHs w/ Pub Assist Income $\leq$ 2%	6.71	**10.36**	2.01	0.38	1.77	**18.58**
Single-Headed HHs $\leq$ 25%	0.70	8.24	3.01	8.68	0.49	**13.52**
Neighborhood is safe at night	1.26	4.89	**11.66**	**11.97**	**18.34**	**44.65**
Police come when called	1.38	9.82	4.78	**21.50**	8.62	**38.97**

Table 4: F-Statistics: Adults
2000 Census Tract Characteristics, Tract of Residence in 2001

Continuous Treatment	Baltimore	Boston	Chicago	LA	NYC	All Sites
Neighborhood Characteristics						
Percent in poverty	7.41	**36.18**	4.74	**54.23**	**34.99**	**101.40**
Percent minority	2.62	5.95	3.88	**10.36**	3.24	**15.93**
Female HS Grad Rate	8.53	**25.94**	3.59	**31.71**	**32.50**	**75.41**
Female BA Attain Rate	6.38	5.10	6.57	**23.15**	**21.20**	**50.12**
Female LFP	**13.67**	**10.59**	2.00	**46.82**	**25.78**	**61.73**
Female High Stat Occ	6.39	6.27	7.06	**29.97**	**13.04**	**47.93**
HHs w/ Pub Assist Income	3.39	**21.00**	2.82	**46.52**	**25.35**	**63.67**
Single-Headed HHs	5.11	**20.51**	3.26	**27.73**	**14.56**	**45.90**

Table 5: F-Statistics: Youth

Treatment	Baltimore	Boston	Chicago	LA	NYC	All Sites
Program						
Moving w/ MTO voucher	**189.1**	**150.1**	**186.5**	**469.4**	**199.0**	**1,033.3**
Neighborhood Poverty						
Percent in poverty $\leq$20%	0.33	**14.42**	**20.08**	**25.87**	**22.82**	**71.82**
Percent in poverty $\leq$10%	3.03	**16.03**	**30.24**	8.89	3.82	**41.13**
Percent in poverty	1.46	**30.78**	**14.06**	**52.06**	**25.24**	**95.31**
School Characteristics						
Rank$\geq$50th percentile on exams	0.88	3.38	4.69	3.19	0.02	8.00
School Ranking on exams	**14.30**	4.91	**13.36**	**22.10**	6.49	**40.16**
Student/Teacher ratio	2.48	0.03	0.12	1.27	2.07	0.25
Hrs/Week spent on hmwk	0.03	1.56	0.05	1.72	0.82	2.41
Student feels safe at school	0.65	0.56	0.00	5.73	4.23	0.39
School has discipline problems	0.66	0.18	0.46	0.36	0.02	0.00
School has disruption problems	0.02	8.96	1.16	2.69	7.44	1.43
School has good climate	3.79	2.69	0.04	1.98	0.01	0.60
Teacher cares	1.60	2.09	1.46	2.01	0.46	0.19
Peer Group						
Friends in old neighborhood	0.77	5.64	0.22	2.13	**14.13**	**14.25**
Peer group uses drugs	2.69	1.20	0.04	0.82	0.41	2.24
Peer group carries weapon	1.97	0.76	0.01	0.41	0.23	0.59
Neighborhood Safety (Parent Reported)						
Neighborhood is safe at night	0.00	3.94	**13.54**	**24.94**	7.14	**39.74**
Police come when called	1.26	6.59	2.57	**28.80**	3.56	**29.52**
Collective Efficacy (Parent Reported)						
Intervene for skipping school	0.01	7.09	5.31	**25.99**	1.54	**26.72**
Intervene for graffiti	0.15	4.55	4.97	**26.22**	8.08	**34.47**

Table 6: Estimates of Local Average Treatment Effects on Adult Outcomes
Under Various Definitions of Treatment

	Poverty Rate $\leq$ 20%			Poverty Rate $\leq$ 10%			Nbd Safe at Night			Female LFP $\geq$ 57%		
	$\widehat{\beta}^{LATE}$	SE	p	$\widehat{\beta}^{LATE}$	SE	p	$\widehat{\beta}^{LATE}$	SE	p	$\widehat{\beta}^{LATE}$	SE	p
Employment												
Not in Labor Force	–0.26	(0.11)	0.02	–0.57	(0.25)	0.02	–0.36	(0.16)	0.03	–0.34	(0.17)	0.05
Employed	0.10	(0.12)	0.37	0.26	(0.25)	0.29	0.28	(0.16)	0.08	0.29	(0.17)	0.09
Unemployed	0.16	(0.07)	0.03	0.31	(0.15)	0.04	0.08	(0.10)	0.44	0.04	(0.10)	0.71
Income												
HH Head's Income	2,548	(2,276)	0.26	6,217	(4,803)	0.20	6,156	(3,140)	0.05	3,247	(3,311)	0.33
HH Total Income	1,913	(3,632)	0.60	6,254	(8,165)	0.44	4,521	(5,107)	0.38	2,239	(5,724)	0.70
Earnings in 2001	1,677	(2,596)	0.52	3,391	(5,544)	0.54	4,743	(3,281)	0.15	2,005	(3,794)	0.60
Welfare Benefits	–0.05	(0.11)	0.66	–0.15	(0.25)	0.54	–0.06	(0.16)	0.71	–0.10	(0.18)	0.57
Health												
Mental Distress	–0.20	(0.08)	0.01	–0.38	(0.18)	0.03	–0.18	(0.11)	0.10	–0.16	(0.12)	0.18
BMI	–4.69	(1.82)	0.01	–7.63	(3.42)	0.03	–5.44	(2.61)	0.04	–3.80	(2.52)	0.13
Depression Last 12 Mos	–0.19	(0.09)	0.03	–0.38	(0.20)	0.06	–0.21	(0.12)	0.09	–0.13	(0.14)	0.33
1st Stage Regression												
F-Statistic	107.3			46.19			30.6			32.7		
Included Sites	Boston, Chicago, LA, NYC			Boston, LA, NYC			Chicago, LA, NYC			Baltimore, LA, NYC		
P(Z=1)	0.363			0.154			0.679			0.271		
P(Z=0)	0.145			0.035			0.498			0.092		

Table 7: Estimates of Local Average Treatment Effects on Youth Outcomes
Under Various Definitions of Treatment

	Poverty Rate $\leq$ 20%			Poverty Rate $\leq$ 10%			Nbd Safe at Night		
	$\widehat{\beta}^{LATE}$	SE	p	$\widehat{\beta}^{LATE}$	SE	p	$\widehat{\beta}^{LATE}$	SE	p
Problem Behaviors									
Ever Smoked Cigarette	0.10	(0.12)	0.38	0.30	(0.36)	0.40	0.22	(0.18)	0.22
Ever Drunk Alcohol	–0.16	(0.13)	0.24	–0.26	(0.37)	0.49	0.04	(0.17)	0.82
Ever Arrested	–0.08	(0.11)	0.48	–0.13	(0.34)	0.70	–0.11	(0.19)	0.54
School									
Currently Enrolled in School	–0.01	(0.13)	0.97	–0.14	(0.37)	0.70	0.07	(0.19)	0.72
Idle	–0.07	(0.10)	0.52	–0.08	(0.31)	0.81	–0.21	(0.17)	0.21
Health									
Asthma	0.13	(0.10)	0.23	0.31	(0.29)	0.28	0.00	(0.14)	0.99
BMI Percentile	1.81	(8.13)	0.82	9.27	(22.03)	0.67	–1.27	(12.16)	0.92
Ever Depressed	–0.06	(0.06)	0.28	0.04	(0.18)	0.82	–0.13	(0.10)	0.17
Mental Distress	–0.09	(0.07)	0.23	–0.14	(0.19)	0.45	–0.18	(0.10)	0.09
1st Stage Regression									
F-Statistic	67.87			30.20			23.68		
Included Sites	Boston, Chicago, LA, NYC			Boston, Chicago			Chicago, LA		
P(Z=1)	0.348			0.133			0.697		
P(Z=0)	0.144			0.030			0.502		

Table 8: Estimates of Local Average Treatment Effects on Youth Outcomes
Treatment is Poverty Rate $\leq$ 20%

	Male			Female		
	$\widehat{\beta}^{LATE}$	SE	p	$\widehat{\beta}^{LATE}$	SE	p
Problem Behaviors						
Ever Smoked Cigarette	0.56	(0.29)	0.05	–0.17	(0.14)	0.22
Ever Drunk Alcohol	0.31	(0.25)	0.21	–0.16	(0.17)	0.37
Ever Arrested	0.00	(0.27)	0.99	–0.07	(0.08)	0.37
School						
Currently Enrolled in School	–0.24	(0.25)	0.34	0.05	(0.15)	0.73
Idle	0.04	(0.21)	0.83	–0.17	(0.11)	0.14
Health						
Asthma	–0.03	(0.18)	0.87	0.10	(0.17)	0.57
BMI Percentile	17.92	17.43	0.30	–15.85	(11.48)	0.17
Ever Depressed	–0.01	(0.08)	0.89	–0.28	(0.12)	0.02
Mental Distress	–0.09	(0.14)	0.54	–0.23	(0.11)	0.05
1st Stage Regression						
F-Statistic	16.70			39.16		
Included Sites	Boston, Chicago, LA			LA, NYC		
P(Z=1)	0.364			0.327		
P(Z=0)	0.184			0.039		

Table 9: Percent of City's Minority Group that lived in Neighborhoods 20% White or Less in 1990

City	Black	Other
All MTO Sites Together	57.4	11.3
Baltimore	56.5	4.7
Boston	37.6	8.1
Chicago	70.8	4.2
Los Angeles	44.9	9.5
New York	57.3	17.8

Table 10: Estimates of ITT Effects on Outcomes, Control Means, and Baseline Means by Site

	Baltimore		Boston		Chicago		LA		NYC	
	$\widehat{\beta}^{ITT}$ or $\widehat{\mu}$	SE	$\widehat{\beta}^{ITT}$ or $\widehat{\mu}$	SE	$\widehat{\beta}^{ITT}$ or $\widehat{\mu}$	SE	$\widehat{\beta}^{ITT}$ or $\widehat{\mu}$	SE	$\widehat{\beta}^{ITT}$ or $\widehat{\mu}$	SE
Adult Outcomes										
HH Head's Income	–1,344	(910)	–57	(914)	6	(807)	–56	(874)	2,861	(845)
Control Mean	12,696	(688)	14,047	(694)	11,075	(618)	11,783	(655)	9,905	(637)
Employed	0.02	(0.05)	–0.04	(0.04)	0.01	(0.04)	0.01	(0.05)	0.11	(0.04)
Control Mean	0.58	(.04)	0.59	(0.03)	0.54	(0.03)	0.46	(0.04)	0.45	(0.03)
Not in Labor Force	0.01	(0.05)	–0.02	(0.04)	–0.03	(0.04)	–0.06	(0.05)	–0.11	(0.04)
Control Mean	0.30	(0.04)	0.36	(0.03)	0.32	(0.03)	0.46	(0.04)	0.45	(0.03)
Youth Outcomes										
Ever Arrested (Males)	0.09	(0.08)	–0.01	(0.06)	0.02	(0.06)	–0.05	(0.06)	0.09	(0.05)
Control Mean	0.30	(0.06)	0.15	(0.04)	0.33	(0.05)	0.28	(0.05)	0.08	(0.04)
Ever Arrested (Females)	0.06	(0.07)	–0.05	(0.04)	–0.01	(0.04)	–0.03	(0.03)	–0.02	(0.03)
Control Mean	0.19	(0.05)	0.10	(0.03)	0.11	(0.03)	0.07	(0.02)	0.05	(0.02)
Ever Smoke Cigarette (Females)	–0.15	(0.06)	–0.10	(0.06)	0.04	(0.05)	0.01	(0.05)	–0.04	(0.05)
Control Mean	0.29	(0.05)	0.32	(0.04)	0.13	(0.04)	0.12	(0.03)	0.20	(0.04)
Baseline Characteristics										
Share Movers	0.53		0.44		0.31		0.66		0.44	
Share Black	0.97		0.37		0.99		0.49		0.48	
Share Hispanic	0.02		0.47		0.01		0.47		0.50	
Share HOPE VI	0.46		0.20		0.36		0.15		0.00	
Car Ownership	0.05		0.22		0.13		0.40		0.06	

Table 11: Estimates of ITT Effects on Neighborhood Characteristics and Control Means by Site

	Baltimore		Boston		Chicago		LA		NYC	
	$\widehat{\beta}^{ITT}$ or $\widehat{\mu}$	SE	$\widehat{\beta}^{ITT}$ or $\widehat{\mu}$	SE	$\widehat{\beta}^{ITT}$ or $\widehat{\mu}$	SE	$\widehat{\beta}^{ITT}$ or $\widehat{\mu}$	SE	$\widehat{\beta}^{ITT}$ or $\widehat{\mu}$	SE
Neighborhood Characteristics										
Streets Safe at Night	0.06	(0.05)	0.11	(0.04)	0.15	(0.04)	0.16	(0.05)	0.21	(0.04)
Control Mean	0.63	(0.04)	0.62	(0.03)	0.60	(0.03)	0.44	(0.04)	0.45	(0.03)
Percent Minority	–4.46	(2.51)	–8.48	(2.71)	–2.42	(1.28)	–5.87	(1.63)	–3.57	(1.48)
Control Mean	87.85	(1.86)	72.38	(2.06)	95.97	(0.98)	93.34	(1.21)	93.89	(1.12)
Share Single-Headed HHs	–0.05	(0.02)	–0.08	(0.02)	–0.03	(0.02)	–0.08	(0.02)	–0.08	(0.02)
Control Mean	0.72	(0.02)	0.60	(0.01)	0.71	(0.01)	0.49	(0.01)	0.66	(0.01)
Share of Females HS Grads	0.05	(0.02)	0.06	(0.01)	0.03	(0.01)	0.11	(0.02)	0.09	(0.01)
Control Mean	0.59	(0.01)	0.66	(0.01)	0.65	(0.01)	0.41	(0.01)	0.52	(0.01)
Share of Females Hold BA	0.03	(0.01)	0.02	(0.01)	0.02	(0.01)	0.04	(0.01)	0.05	(0.01)
Control Mean	0.11	(0.01)	0.15	(0.01)	0.10	(0.01)	0.06	(0.01)	0.09	(0.01)
Female LFP Rate	0.05	(0.01)	0.03	(0.01)	0.01	(0.01)	0.06	(0.01)	0.06	(0.01)
Control Mean	0.49	(0.01)	0.53	(0.01)	0.52	(0.01)	0.43	(0.01)	0.40	(0.01)
School Characteristics										
School Percentile (Test Scores)	5.87	(1.41)	3.55	(1.32)	2.73	(0.69)	5.93	(1.08)	3.67	(1.45)
Control Mean	15.55	(1.05)	19.25	(0.98)	9.27	(0.53)	16.33	(0.81)	13.49	(1.07)