working paper

11 01

Assessing the Evidence on Neighborhood Effects from Moving to Opportunity

by Dionissi Aliprantis

FEDERAL RESERVE BANK OF CLEVELAND

Working Paper 11-01 January 2011

Assessing the Evidence on Neighborhood Effects from Moving to Opportunity

by Dionissi Aliprantis

The interpretation of estimates from Moving to Opportunity (MTO) as neighborhood effects has created significant controversy among social scientists. This paper presents a framework that clarifies the interpretation of results from the MTO housing mobility experiment. The paper defines several neighborhood treatments and estimates their Local Average Treatment Effects (LATEs) using assigned treatment in MTO as an instrumental variable. This framework clarifies that while parameters estimated in the literature do not suffer from selection bias, selection into treatment is an inescapable issue if one seeks to learn about neighborhood effects from MTO. The LATE parameters estimated in this paper are neighborhood effects for the subgroup of MTO families who are compliers with respect to the defined treatment. In contrast, the Treatment-on-the-Treated (TOT) parameters reported in the literature are program effects. Since the subgroup of compliers for various neighborhood treatments can be considerably smaller than the subgroup induced to move by MTO, preliminary estimates indicate that LATE neighborhood effects tend to be much larger than the TOT program effects from MTO. This re-interpretation of the MTO data suggests two important conclusions related to the current understanding of neighborhood effects and programs. First, if alternative housing mobility programs were designed to induce moves to neighborhoods with characteristics other than low poverty, it is entirely feasible that such programs might induce larger effects than MTO. Second, initial LATE estimates appear to reconcile the evidence from MTO with prevailing theories of neighborhood effects.

Key words: 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 code: C30, H50, I38, J10, R00.

Dionissi Aliprantis is at the Federal Reserve Bank of Cleveland. He can be reached at Research Department, P.O. Box 6387, Cleveland, OH 44101-1387, USA; (216)579-3021; or dionissi.aliprantis@clev.frb.org. The author thanks Francisca G.-C. Richter for helpful comments and Mary Zenker for valuable research assistance. 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. The views stated here are those of the author and are not necessarily those of the U.S. Department of Education.

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. The gaps between black and white children in educational outcomes such as test scores (Reardon (2008)) and attainment (Heckman and LaFontaine (2010)) are well documented. These gaps are associated with later outcomes of importance: early skill differences are able to explain a large share of the subsequent racial earnings gap (Neal and Johnson (1996), Keane and Wolpin (2000)), and Pettit and Western (2004) estimate that for black (white) males aged 30-34 in 1999, nearly 59% (11%) of high school dropouts had spent time in prison. African Americans are also exposed to much higher levels of violence than their white counterparts. In 2006 the homicide death rate of black males between 15-34 was approximately 8 times that of white males (NCHS (2009)), as it had been for years (Figure 1).

Although these outcomes have received much attention from social scientists, the mechanisms maintaining racial gaps are not well understood. One prominent theory proposes that neighborhood effects can explain these differences in outcomes. Wilson (1987) presents empirical evidence that urban poverty in the US, in particular that of predominantly African American neighborhoods, has become highly concentrated over recent decades. For example, 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 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 appear to reproduce the beneficial effects found in Gautreaux.

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. This paper uses assigned treatment in MTO as an instrumental variable to estimate the LATEs of various neighborhood treatments induced by changes in this instrument. The neighbor-

[1] However, this number dropped in the 1990s (Jargowsky (2003)).

[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)).

hood treatments induced by MTO are ascertained by comparing the neighborhood characteristics of the MTO groups, and a key lesson from this exercise is that alternative housing mobility programs might result in different neighborhood effects than those observed in MTO. That is, the LATE is defined by the subgroup of compliers, and thus different instruments will result in different LATE parameters if they induce different subpopulations to select into treatment (Heckman (1997)).

Estimating LATEs from MTO helps to clarify that while parameters estimated in the literature do not suffer from selection bias (Ludwig et al. (2008)), selection into treatment is an inescapable issue if one seeks to learn about neighborhood effects from MTO (Clampet-Lundquist and Massey (2008)). Preliminary estimates indicate that many LATE neighborhood effects will be larger than the TOT program effects from MTO. For example, it is shown that moving to a neighborhood with a high degree of personal safety, moving to an integrated neighborhood, and moving to a neighborhood with a high level of collective efficacy through MTO all have much larger effects on adult outcomes than moving to a low-poverty neighborhood through MTO. Because LATE parameters are defined by the subpopulation of compliers, these neighborhood effects pertain to much smaller shares of MTO households than the program effects reported in the literature, since the assigned treatment of housing vouchers only induced a small fraction of MTO households to receive these neighborhood treatments. While the current analysis is only able to estimate a subset of LATE parameters of interest using publicly available information on MTO, future analyses to be conducted on restricted access data will illustrate these issues further.

This re-interpretation of the MTO data suggests two important conclusions related to the current understanding of neighborhood effects and programs. First, if alternative housing mobility programs were designed to induce moves to neighborhoods with characteristics other than low poverty, it is entirely feasible that such programs might induce larger effects than MTO. Second, initial LATE estimates appear to reconcile the evidence from MTO with prevailing theories of neighborhood effects (Wilson (1987)).

The paper proceeds as follows: Section 2 describes the MTO experiment and presents some descriptive statistics. Section 3 draws heavily from the line of research summarized in Heckman (2010) and Heckman and Vytlacil (2005) to define and interpret LATE parameters when assigned treatment is viewed as an instrumental variable. A summary of the program effects found in the literature is presented in Section 4. Section 5 first discusses the implications of the LATE for the interpretation of neighborhood effect parameters. Section 5 next defines several neighborhood treatments and then presents estimates of their LATEs resulting from MTO. Section 6 concludes.

2 Moving To Opportunity (MTO)

Moving To Opportunity (MTO) was inspired by the promising results of the Gautreaux program. At the same time that much attention was being devoted to the increasing concentration of poverty in the US (Wilson (1987)), results from the Gautreaux program indicated that housing mobility could be an effective policy to mitigate the adverse effects of concentrated poverty.

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. The resulting program created by HUD and CHA gave families awarded Section 8 public housing vouchers the ability to use them beyond the territory of CHA.

Specifically, the Gautreaux court ruling allowed families to be relocated either to suburbs that were at least 70 percent white 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, 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 decentralize 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 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

[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 the common cutoff considered in the social sciences - census tracts with 40% or more of residents poor, while the threshold of low poverty was set at the median tract-level poverty rate in 1990 - 10% (Goering (2003)).

then unrestricted in where they used their Section 8 vouchers. Families in this group were also provided with counseling and education [on housing markets/home finances?] 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. More information on MTO may be found on HUD's MTO webpage or the following online repository of papers on MTO.

2.1 Data and Descriptive Statistics

Due to the detailed geographic information they contain, the data from MTO are neither publicly nor privately available. As a result all of the data used in the current paper have been collected from the literature. The current results should be viewed as an interim analysis, as a request for access to the MTO Restricted Access Data has been made by the author.

There were 4,248 (4,610) accepted (applicant) families, with 1,310 (1,440) families in the control group, 1,209 (1,350) families in the Section 8 only group, and 1,729 (1,820) 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% 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))

3 The Identification of Treatment Effects in Social Experiments

Before interpreting results from MTO, we first define and state identifying assumptions for the treatment effects found in the MTO literature. 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 Assigned Treatment and Selection into Treatment

Since the parameters of interest in evaluating programs are often defined in terms of the subpopulation receiving treatment, we model how individuals select into treatment. We begin by noting that in the case of social experiments, a researcher can typically control assignment but not receipt of treatment. Define Z as an indicator for the treatment assigned to an individual so that:

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

Since it need not be true that $D = Z$, we write $D(Z)$ to denote the treatment received when assigned treatment Z. Furthermore, suppose there is a latent index D^* that depends on assigned treatment Z and some unobserved component U as follows:

$$D^* = \mu(Z) - U, \tag{2}$$

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

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

$$D = \begin{cases} 1 & \text{if } D^* \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

We follow Heckman and Vytlacil (2000) and assume:

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

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

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 2 and 3 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 $U = F_V(V(Z))$. We write the propensity score as $P(z) = Pr(D = 1|Z = z)$. Note that Assumption 4 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 2 and 3, 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, $U \sim U[0, 1]$, then 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.[6]

Note that if $D^* = \mu_1(Z) - U_1$ for all individuals as in Figure 2a, then there are no defiers.[7] 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

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

[7] In order for the selection model in Equations 2 and 3 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 (2010)) 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.

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. Determining 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.

3.2 The ITT, LATE, and TOT Parameters

Given our joint model of outcomes (Equation 1) and selection into treatment (Equations 2 and 3), we now consider the assumptions necessary to identify parameters of interest. We will use this joint model to define and interpret these parameters. We begin by comparing the outcome variable Y at two different values of treatment assignment, $Z = 1$ and $Z = 0$, to obtain the Wald estimator:

$$\begin{aligned}
&E[Y|Z = 1] - E[Y|Z = 0] \\
&\quad = E[D(1)Y(1) + (1 - D(1))Y(0)|Z = 1] - E[D(0)Y(1) + (1 - D(0))Y(0)|Z = 0] \\
&\quad = E[(D(1) - D(0))(Y(1) - Y(0))] &(4)\\
&\quad = Pr[D(1) - D(0) = 1]\, E[Y(1) - Y(0)|D(1) - D(0) = 1] &(5)\\
&\qquad + Pr[D(1) - D(0) = -1]\, E[Y(0) - Y(1)|D(1) - D(0) = -1].
\end{aligned}$$

Equation 4 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:

$$\beta^{ITT} \equiv E[Y|Z = 1] - E[Y|Z = 0]. \tag{6}$$

The only assumptions necessary to identify the ITT effect from the Wald estimator in Equation 5 are Assumptions 1 and 2. Under these assumptions, Equation 5 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.

Researchers are usually 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 5 to identify treatment effects that go beyond the ITT and inform us about the effect of treatment on outcomes.[8] Much of the literature on instrumental variables does this by placing restrictions on how changes

[8] 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*.

in the instrument induce changes in treatment (ie, on the selection model in Equations 2 and 3.). 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 2 and 3 that allow for the identification of treatment effects when combined with Assumptions 1–4. For example, 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 5 identifies

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

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. That is, the LATE informs us about the average effect of treatment on compliers.

Alternatively, the researcher may believe that 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

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

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. Note that Assumption 6 implies Assumption 5, but Assumption 5 does not imply Assumption 6. Assumption 6 is the special case of Assumption 5 in which there are no always-takers, and note that under Assumption 6 the LATE and TOT parameters will coincide:

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

Alternatively, 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 2 and 3, together with Assumptions 1-5, the following parameters exist and are finite:

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

As stated earlier, when $\mu_1(Z=0) \leq 0$, then these two parameters are the same. 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).

4 Program Effects

The ITT effects identified in the literature on MTO compare the 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 (Rosenbaum (1995)), 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.[9] 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 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). These sites were almost entirely African-American (unlike the other sites, which also had some Hispanic households), and also had higher crime rates (Burdick-Will et al. (2009)).

5 Neighborhood Effects

There is considerable controversy surrounding the interpretation of results from MTO as neighborhood effects. One interpretation of results argues that selection into treatment biases parameter estimates (Clampet-Lundquist and Massey (2008)). An opposing interpretation is that randomiza-

[9] 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).

tion 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."[10]

How can we reconcile these opposing views? We begin by noting 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). Randomization occurred in MTO at the level of assigned treatment ($Z \in \{0, 1\}$), not at the level of treatment ($D \in \{0, 1\}$). Thus households were able to choose whether or not to move after receiving their assigned treatment, and these choices were not made randomly. It is for this reason that we can expect variation in neighborhood of residence *between* compliers, always-takers, and never-takers to be correlated with individual characteristics.[11] This was precisely the case in MTO: 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), p 86).

Thinking about variation in neighborhood of residence by the subgroups in Table 1 is helpful because it clarifies that we can only expect variation in neighborhood of residence induced by randomization to be uncorrelated with individual characteristics *within* subgroups. But since there should be no experimentally induced variation in neighborhood of residence within the group of always-takers, and the same should be true within the group of never-takers, housing vouchers offered through MTO should only have induced variation in neighborhood of residence for the subgroup of compliers. Thus, while it is true that randomization in MTO did indeed induce variation in neighborhood of residence that was uncorrelated with individual characteristics, this statement is only true *within* the subgroup of compliers.

What are the consequences of exogenous variation in neighborhood of residence being restricted to the subgroup of compliers? On the one hand, this point has no implications for reinterpreting estimates of program effects 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 so TOT and LATE parameters are identical. These TOT effects are all qualitatively similar to the ITT effects just reviewed, but with 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

[10]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).

[11]Remember that we assume there are no defiers.

effect of a 5 percentage point decrease (Kling et al. (2007a)).

On the other hand, careful consideration of selection into treatment is likely to have major implications for the interpretation of evidence from MTO if we are interested in understanding neighborhood effects. If, for example, 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 the same. Looking at Equations 7 and 8, in this case the LATE parameter would no longer be the ITT effect normalized by the percent of people moving, but rather would become the ITT effect normalized by the percent of households induced into treatment by the instrument. This distinction will have major implications, as it allows for the possibility of dramatically different program (TOT) and neighborhood (LATE) effects. The remainder of this section will provide examples and estimates of LATEs to clarify that while parameters estimated in the literature do not suffer from selection bias (Ludwig et al. (2008)), selection into treatment is an inescapable issue if one seeks to learn about neighborhood effects from MTO (Clampet-Lundquist and Massey (2008)).

5.1 Selection into Treatment and LATE Parameters

We briefly consider an example to help illustrate how selection into treatment drives LATE estimates. Define treatment as moving through the MTO program, and suppose that the ITT effect on log wages is 0.030. According to Table F9 in Kling et al. (2007b), $E[D|Z=1]-E[D|Z=0]=0.467$ for the experimental group under this definition of treatment. Thus

$$\beta^{LATE}(Z=1,Z=0)=\frac{E[Y|Z=1]-E[Y|Z=0]}{E[D|Z=1]-E[D|Z=0]}=\frac{0.030}{0.467}=0.064.$$

In this hypothetical example, the LATE estimates that moving through the MTO program causally increased wages by 6.4%.

Now suppose that treatment is defined as moving to an integrated neighborhood (ie, a neighborhood in which at least 50% of residents are white). According to Table F2 of Kling et al. (2007b), $E[D|Z=1]-E[D|Z=0]=0.065$ for the experimental group under this definition of treatment. This will yield:

$$\beta^{LATE}(Z=1,Z=0)=\frac{E[Y|Z=1]-E[Y|Z=0]}{E[D|Z=1]-E[D|Z=0]}=\frac{0.030}{0.065}=0.462.$$

Under this hypothetical case, moves to integrated neighborhoods induced by MTO would have causally increased wages by 46%! Note that not only has the LATE point estimate increased dramatically here, but the subgroup of compliers to which this estimate applies has shrunk dramatically.

To further illustrate the importance of selection into treatment for interpreting LATE parameters, now suppose that we observe a LATE for a given definition of treatment. Different effects and patterns of selection into treatment on the part of compliers and always-takers could generate this parameter value. Define treatment as "moving to a neighborhood with a poverty rate of no more than 20%," and assume that 50% of those assigned treatment (ie, were given a housing voucher) selected into treatment. Let Y now denote change in log wages, and suppose that moving to a neighborhood with less than 20% poverty increases both compliers' and always-takers' wages by 10%. Then we might have the following scenario if 40% are compliers:

$$\beta^{LATE}(Z=1, Z=0) = \frac{[.4 \cdot 0.10 + .1 \cdot 0.10 + .5 \cdot 0] - [.1 \cdot 0.10 + .9 \cdot 0]}{.5 - .1} = 0.10.$$

Note that the LATE is not informative about the effect on always-takers' wages. If compliers' wages increased by 10% but always-takers' wages increased by 20%, then the LATE would still be:

$$\beta^{LATE}(Z=1, Z=0) = \frac{[.4 \cdot 0.10 + .1 \cdot 0.20 + .5 \cdot 0] - [.1 \cdot 0.20 + .9 \cdot 0]}{.5 - .1} = 0.10.$$

As well, note that if the group of compliers shrinks dramatically, the interpretation of the LATE changes while its value stays the same. Suppose that now 40% of the population are always-takers, and only 10% are compliers. In this case the LATE is:

$$\beta^{LATE}(Z=1, Z=0) = \frac{[.1 \cdot 0.10 + .4 \cdot 0.20 + .5 \cdot 0] - [.4 \cdot 0.20 + .6 \cdot 0]}{.5 - .4} = 0.10.$$

These hypothetical examples help to illustrate two crucial points about LATE parameters. First, LATE estimates can only be understood after defining treatment and determining the share of the population induced to select into this treatment by the instrument. Second, 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 in Gautreaux were compliers (Ludwig et al. (2008)), its alternative rules for moving likely induced a greater share of compliers for many neighborhood treatments than MTO. Thus the LATE parameters defined by MTO are only a subset of possible neighborhood effects, even amongst those resulting from housing mobility programs. Different housing mobility programs designed to induce moves to different types of neighborhoods could plausibly result in different neighborhood effects.

5.2 Preliminary LATE Estimates for Alternative Definitions of Treatment

To learn about neighborhood effects from MTO, we now present various definitions of treatment. Since the threshold for low poverty in the design of MTO was set at the median Census tract level poverty rate in 1990, 10% (Goering (2003)), reasonable cutoffs for defining neighborhood treatments would be families moving to a neighborhood at or above the national median for a given characteristic in 1990. Table 2 shows the median US census tract characteristics in 1990 and 2000 for several neighborhood characteristics that one might believe affect outcomes.

None of the LATE parameters defined with respect to the median US Census tract characteristics as shown in Table 2 may be estimated with publicly available data from MTO. Thus this paper is only able to present a preliminary analysis of LATE estimates. As mentioned in Section 2.1, the author has submitted a request for access to the MTO Restricted Access Data. A full analysis will be performed upon receipt of the Restricted Access Data.

The LATE parameters reported in this analysis are those on adult outcomes for the experimental group. The estimates are constructed by comparing ITT effects using the Wald estimator from Equation 7:

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

All estimates are taken from Kling et al. (2007b).[12]

5.2.1 Moving to a Low-Poverty Neighborhood

Define treatment as moving to a census tract in which the poverty rate was below 30%. Then according to Table F2 in Kling et al. (2007b), $E[D|Z=1] - E[D|Z=0] = 0.345$ for the experimental group. Note that by using 30% poverty as a cutoff for treatment, under this definition many more families will select into treatment than under the definition using the 1990 median census tract poverty rate of 10%. Table 3 shows LATE estimates for neighborhood effects on adult outcomes. This is one neighborhood effect about which we have learned much through MTO. LATE parameters of this neighborhood effect do not look tremendously different than the TOT parameters found in the literature; such moves do not have large effects on adult outcomes.

5.2.2 Moving to a Neighborhood with a High Degree of Personal Safety

Now suppose that we are interested in how personal safety affects adult outcomes (Anderson (1999)). We might define treatment either as "moving to a neighborhood in which one has not seen illicit drugs being sold or used during the past 30 days," or alternatively as "moving to a neighborhood in which the residents believe streets are safe or very safe at night." According to Table F9 of Kling et al. (2007b), $E[D|Z=1] - E[D|Z=0] = 0.118$ under the first definition and $E[D|Z=1] - E[D|Z=0] = 0.141$ under the second definition. Tables 4 and 5 show LATEs for these neighborhood treatments. Since we are normalizing the original ITT effects by a much smaller share of compliers, the LATE parameters become much larger than the earlier LATEs of moving to a low-poverty neighborhood. Although not precisely estimated in this preliminary analysis, point estimates suggest that effects from this neighborhood treatment could be large.

[12] Standard errors are calculated using a Taylor approximation for the variance of the ratio of two random variables as presented in Section 5.5 of Casella and Berger (2002):

$$V(\tfrac{X}{Y}) = \left(\frac{\mu_X}{\mu_Y}\right)^2 \left(\frac{V(X)}{\mu_X^2} + \frac{V(Y)}{\mu_Y^2} - 2\frac{Cov(X,Y)}{\mu_X \mu_Y}\right).$$

We assume $Cov(X,Y) = 0$, so these are conservative standard errors.

5.2.3 Moving to an Integrated Neighborhood

Define treatment as moving to a neighborhood with less than 50% of residents being minorities. According to Table F2 of Kling et al. (2007b), $E[D|Z = 1] - E[D|Z = 0] = 0.065$ for the experimental group. Table 6 shows LATEs for this neighborhood treatment. Again, although these LATEs are not precisely estimated in this preliminary analysis, point estimates suggest that effects from this neighborhood treatment could be large.

5.2.4 Moving to a Neighborhood with High Levels of Collective Efficacy

Finally, suppose that we are interested in how neighborhood-level collective efficacy affects adult outcomes (Sampson et al. (1997)). We might define treatment either as "moving to a neighborhood in which residents believe neighbors are likely to intervene against graffiti," or alternatively as "moving to a neighborhood in which residents believe their neighbors are likely to intervene if they encounter kids skipping school." Tables 7 and 8 show LATEs for these neighborhood treatments. A similar pattern is again observed; a smaller share of compliers makes the point estimates of LATEs from this neighborhood treatment much larger than the LATEs from the low-poverty treatment.

5.2.5 Further Neighborhood Treatments

Once restricted access data is obtained, the author will estimate LATE parameters for neighborhood effects defined for the treatments found in Table 2, among others. Some of the LATE parameters to be estimated will address the effects of moving to:

- a neighborhood with a high labor force participation rate
- a high quality school (as measured by both test scores and student/teacher ratios)
- a neighborhood with high educational attainment (as measured by attainment of both bachelor's degrees and high school diplomas)
- a neighborhood with a high rate of two-parent families
- a neighborhood with high wealth

6 Conclusion

Estimating LATEs from MTO helps to clarify that while parameters estimated in the literature do not suffer from selection bias (Ludwig et al. (2008)), selection into treatment is an inescapable issue if one seeks to learn about neighborhood effects from MTO (Clampet-Lundquist and Massey (2008)). Preliminary estimates indicate that many LATE neighborhood effects will be larger than the TOT program effects from MTO. This re-interpretation of the MTO data suggests two important conclusions related to the current understanding of neighborhood effects and programs. First, if alternative housing mobility programs were designed to induce moves to neighborhoods with characteristics other than low poverty, it is entirely feasible that such programs might induce larger effects than MTO. Second, initial LATE estimates appear to reconcile the evidence from MTO with prevailing theories of neighborhood effects (Wilson (1987)).

References

Aliprantis, D. (2010). Redshirting, compulsory schooling laws, and educational attainment. *Journal of Educational and Behavioral Statistics*. Forthcoming.

Anderson, E. (1999). *Code of the Street: Decency, Violence, and the Moral Life of the Inner City.* W. W. Norton and Company.

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.

Burdick-Will, J., J. Ludwig, S. W. Raudenbush, R. J. Sampson, L. Sanbonmatsu, and P. Sharkey (2009). Converging evidence for neighborhood effects on children's test scores: An experimental, quasi-experimental, and observational comparison. *Mimeo., Brookings Institution*.

Casella, G. and R. L. Berger (2002). *Statistical Inference*. Duxbury.

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.

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.

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).

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.

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., 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.

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.

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.

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

Sampson, R. J., S. W. Raudenbush, and F. Earls (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science 277*(15), 918–924.

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.

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

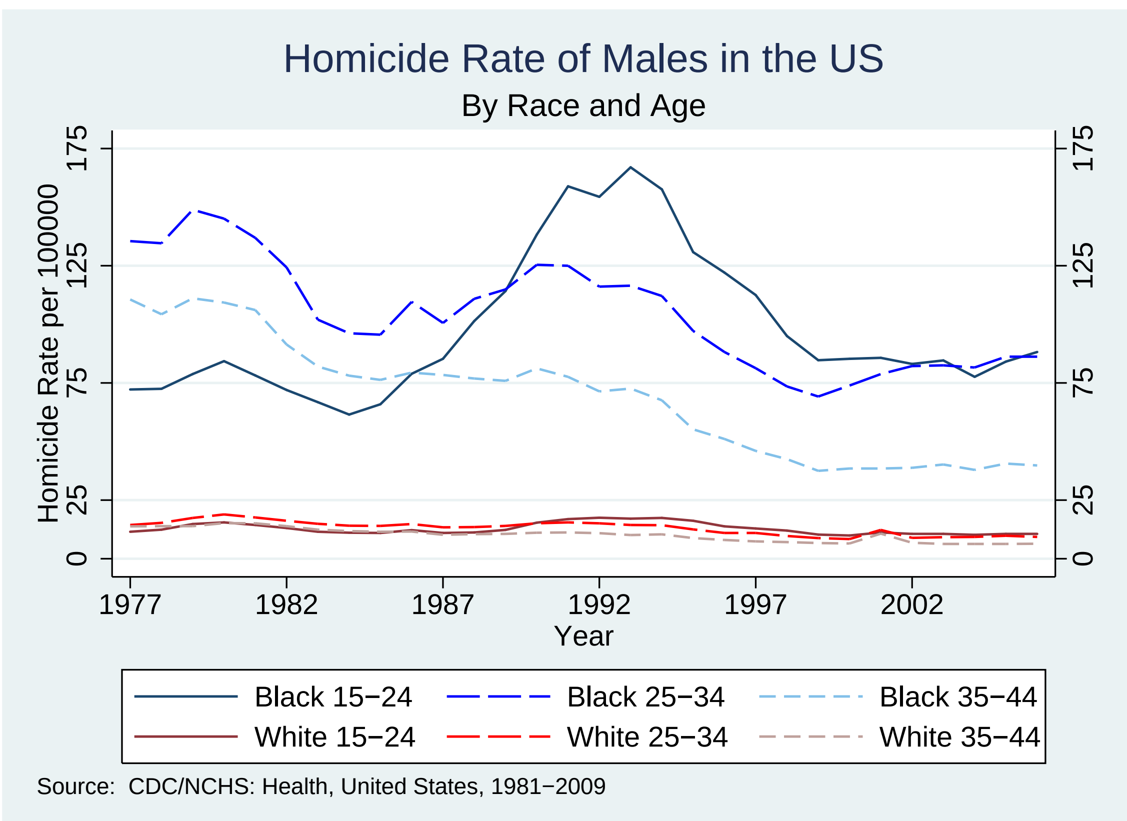


Figure 1: Homicide Rate of Males

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

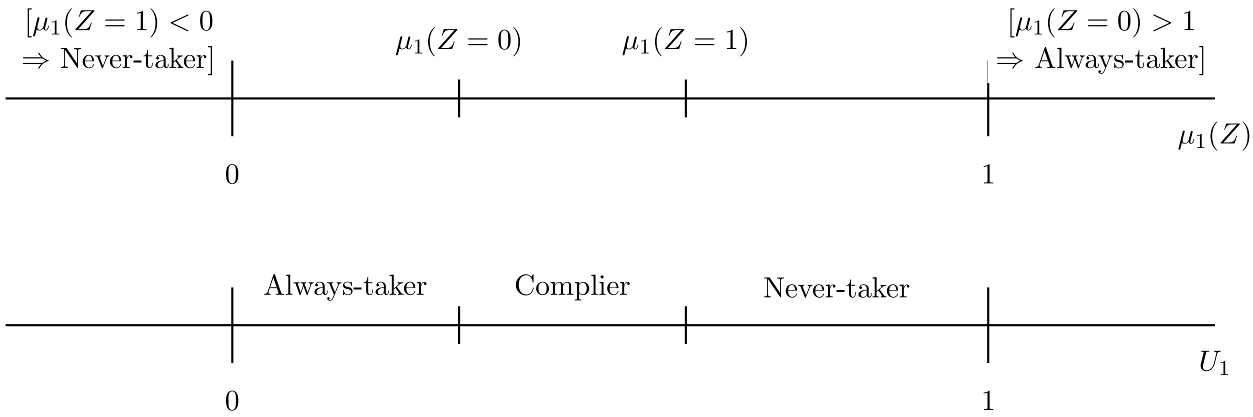
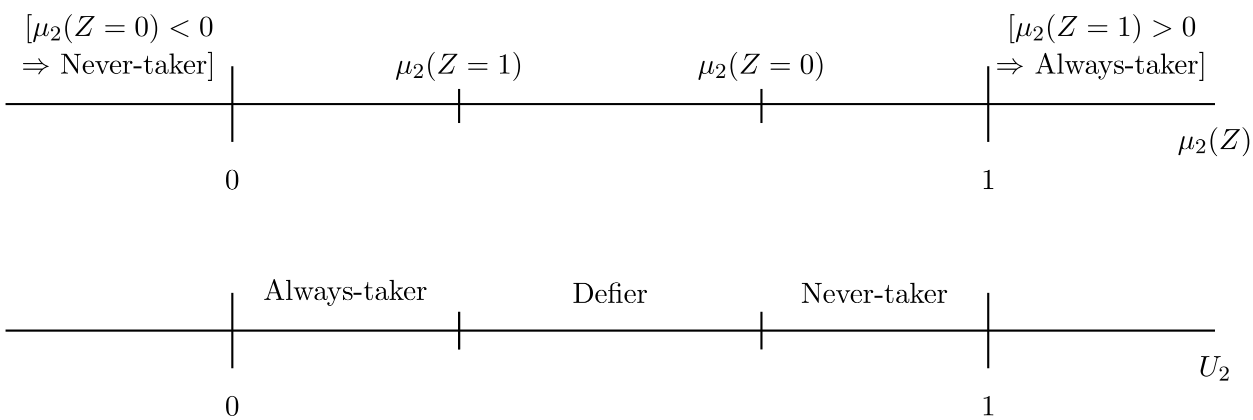


Figure 2b: Type 2 Households: $\mu_2(Z=1) < \mu_2(Z=0)$



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

Figure 2: Selection into Treatment

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: United States Median Census Tract Characteristics in 1990 and 2000

	Median Census Tract 1990	Median Census Tract 2000	Sample
Neighborhood Characteristic (%)			
White	91.69	85.23	Summary File 1
Unemployed (Male)	5.59	4.83	Age≥16, Summary File 3
Unemployed (Female)	5.40	4.80	Age≥16, Summary File 3
Labor Force Participation Rate (Male)	74.95	71.73	Age≥16, Summary File 3
Labor Force Participation Rate (Female)	56.13	57.43	Age≥16, Summary File 3
Residents with ≥HS Diploma	75.61	82.21	Age≥25, Summary File 3
Residents with ≥BA	14.14	17.86	Age≥25, Summary File 3
Two-Parent Families	82.21	72.72	Family Households, Summary File 3

Table 3: Effects of Moving to a Low Poverty Neighborhood ($\leq$ 30% Poverty Rate) due to MTO
(Source of ITT Estimates: Kling et al. (2007b))

Outcome	$\widehat{\beta}^{LATE}$	$E[Y\|Z=1]-E[Y\|Z=0]$ Source	$E[D\|Z=1]-E[D\|Z=0]$ Source
Adult Outcomes			
Employed	0.043	0.015 (0.021)	0.345 (0.018)
	(0.42)	Table F3	Table F2
Earnings in 2001	362	125 (449)	
	(154)	Table F3	
Physical Health	–0.035	–0.012 (0.026)	
Has fair/poor health	(0.47)	Table F6	
Mental Health	–0.261	–0.090 (0.064)	
K6 z-score of psych. stress	(0.74)	Table F6	

Table 4: Effects of Moving to a Safe Neighborhood (Believe Safe at Night) due to MTO
(Source of ITT Estimates: Kling et al. (2007b))

Outcome	$\widehat{\beta}^{LATE}$	$E[Y\|Z=1]-E[Y\|Z=0]$ Source	$E[D\|Z=1]-E[D\|Z=0]$ Source
Adult Outcomes			
Employed	0.106	0.015 (0.021)	0.141 (0.022)
	(1.03)	Table F3	Table F9
Earnings in 2001	887	125 (449)	
	(945)	Table F3	
Physical Health	–0.085	–0.012 (0.026)	
Has fair/poor health	(1.15)	Table F6	
Mental Health	–0.638	–0.090 (0.064)	
K6 z-score of psych. stress	(1.92)	Table F6	

Table 5: Effects of Moving to a Safe Neighborhood (Not Seen Drugs) due to MTO
(Source of ITT Estimates: Kling et al. (2007b))

Outcome	$\widehat{\beta}^{LATE}$	$E[Y\|Z=1]-E[Y\|Z=0]$ Source	$E[D\|Z=1]-E[D\|Z=0]$ Source
Adult Outcomes			
Employed	0.127	0.015 (0.021)	0.118 (0.022)
	(1.24)	Table F3	Table F9
Earnings in 2001	1,059	125 (449)	
	(1,344)	Table F3	
Physical Health	–0.102	–0.012 (0.026)	
Has fair/poor health	(1.37)	Table F6	
Mental Health	–0.763	–0.090 (0.064)	
K6 z-score of psych. stress	(2.35)	Table F6	

Table 6: Effects of Moving to an Integrated Neighborhood ($\geq$ 50% White) due to MTO
(Source of ITT Estimates: Kling et al. (2007b))

Outcome	$\widehat{\beta}^{LATE}$	$E[Y\|Z=1]-E[Y\|Z=0]$ Source	$E[D\|Z=1]-E[D\|Z=0]$ Source
Adult Outcomes			
Employed	0.231	0.015 (0.021)	0.065 (0.015)
	(2.27)	Table F3	Table F2
Earnings in 2001	1,923	125 (449)	
	(3,638)	Table F3	
Physical Health	–0.185	–0.012 (0.026)	
Has fair/poor health	(2.50)	Table F6	
Mental Health	–1.385	–0.090 (0.064)	
K6 z-score of psych. stress	(4.69)	Table F6	

Table 7: Effects of Moving to a Neighborhood with High Collective Efficacy (Neighbors Intervene Against Graffiti) due to MTO (Source of ITT Estimates: Kling et al. (2007b))

Outcome	$\widehat{\beta}^{LATE}$	$E[Y\|Z=1]-E[Y\|Z=0]$ Source	$E[D\|Z=1]-E[D\|Z=0]$ Source
Adult Outcomes			
Employed	0.090	0.015 (0.021)	0.166 (0.049)
	(0.88)	Table F3	Table G5
Earnings in 2001	753	125 (449)	
	(1,012)	Table F3	
Physical Health	–0.072	–0.012 (0.026)	
Has fair/poor health	(0.98)	Table F6	
Mental Health	–0.542	–0.090 (0.064)	
K6 z-score of psych. stress	(1.69)	Table F6	

Table 8: Effects of Moving to a Neighborhood with High Collective Efficacy (Neighbors Intervene if Kids Skipping School) due to MTO (Source of ITT Estimates: Kling et al. (2007b))

Outcome	$\widehat{\beta}^{LATE}$	$E[Y\|Z=1]-E[Y\|Z=0]$ Source	$E[D\|Z=1]-E[D\|Z=0]$ Source
Adult Outcomes			
Employed	0.152	0.015 (0.021)	0.099 (0.053)
	(1.51)	Table F3	Table G5
Earnings in 2001	1,263	125 (449)	
	(2,944)	Table F3	
Physical Health	–0.121	–0.012 (0.026)	
Has fair/poor health	(1.65)	Table F6	
Mental Health	–0.909	–0.090 (0.064)	
K6 z-score of psych. stress	(3.32)	Table F6	