working paper

12 33

Assessing the Evidence on Neighborhood Effects from Moving to Opportunity

Dionissi Aliprantis

FEDERAL RESERVE BANK OF CLEVELAND

Working papers of the Federal Reserve Bank of Cleveland are preliminary materials circulated to stimulate discussion and critical comment on research in progress. They may not have been subject to the formal editorial review accorded official Federal Reserve Bank of Cleveland publications. The views stated herein are those of the authors and are not necessarily those of the Federal Reserve Bank of Cleveland or of the Board of Governors of the Federal Reserve System.

Working papers are available on the Cleveland Fed's website at:

www.clevelandfed.org/research.

Working Paper 12-33 November 2012*

Assessing the Evidence on Neighborhood Effects from Moving to Opportunity

Dionissi Aliprantis

Trying to learn about neighborhood effects from the Moving to Opportunity (MTO) housing mobility experiment by focusing on its program effects obfuscates the evidence on neighborhood effects from MTO. This paper shows that using Intent-to-Treat (ITT) and Treatment-on-the-Treated (TOT) program effects from MTO to indirectly draw conclusions about neighborhood effects (1) offers no advantage for learning about neighborhood effects over directly estimating neighborhood effects, and (2) answers an ill-posed question as a result of allowing central identifying assumptions to be made implicitly. Focusing attention on directly specifying and estimating models of neighborhood effects, the paper presents empirical evidence that MTO only identifies effects from moves between neighborhoods of low quality. These results have broad implications for the way program effects are used to learn about parameters of other models, and they have not been sufficiently addressed in the literature to interpret the evidence from MTO as a test of Wilson (1987).

Keywords: Marginal Treatment Effect, Essential Heterogeneity, Strong Ignorability, Moving to Opportunity, Neighborhood Effect, Concentrated Poverty, Segregation.

JEL Codes: C30, H50, I38, J10, R00.

*Note: This paper is a substantial revision of a paper first released in January 2011 as working paper no. 11-01 and later replaced with working paper no. 11-22. The last version of that paper was released in August 2012, and with the release of this paper (working paper no. 12-32), it has been removed.

Dionissi Aliprantis is at the Federal Reserve Bank of Cleveland (dionissi.aliprantis@clev.frb.org).

1 Introduction

"The problem of the Twentieth Century" has yet to be resolved. The distributions of blacks and whites in the United States are dramatically different for nearly every outcome of importance, and the mechanisms maintaining these differences are not well understood. One prominent theory proposes that effects from living in a poor, segregated, and socially isolated neighborhood can help explain these differences in outcomes (Wilson (1987)). The large differences in the neighborhood environments of blacks and whites (Wilson (1987), Massey and Denton (1993)), as well as the recent increase in the share of Americans living in census tracts with high poverty rates (Jargowsky (1997), Kneebone et al. (2011)), have motivated a large literature to investigate neighborhood effects.

Since households endogenously sort into neighborhoods, researchers have attempted to identify neighborhood effects using the exogenous variation in neighborhoods induced by housing mobility programs. Two of the best known housing mobility programs are the Gautreaux program and the Moving to Opportunity (MTO) housing mobility experiment. The Gautreaux program was designed to desegregate public housing in Chicago and relocated public housing residents through housing vouchers in a quasi-random manner. Those who moved to high-income, white-majority suburbs through Gautreaux had much better education and labor market outcomes than those who moved to segregated city neighborhoods (Rosenbaum (1995), Mendenhall et al. (2006)). MTO was an 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 to be used in low-poverty neighborhoods. In a tremendous disappointment, MTO did not reproduce some of the most beneficial effects found in Gautreaux.

This paper investigates the assumptions under which various parameters can be identified by MTO. A joint model of potential outcomes and selection into treatment is specified, and a variety of identifying assumptions, such as those established in Imbens and Angrist (1994) and Heckman and Vytlacil (2005), are interpreted in the context of MTO. Empirical evidence is presented that MTO only identifies effects from moves between neighborhoods of low quality, regardless of the identifying assumptions adopted.

This paper also shows that the literature has used program effects inappropriately when trying to learn about neighborhood effects from the results of MTO. First, it is shown how the most prominent interpretation of results from MTO in the literature uses Intent-to-Treat (ITT) and Treatment-on-the-Treated (TOT) program effects from MTO to *indirectly* draw conclusions about neighborhood effects. The logic behind this interpretation is expressed formally, illustrating that the indirect approach offers no advantage for learning about neighborhood effects over directly estimating neighborhood effects. Even using this indirect approach, the researcher cannot escape explicitly specifying those neighborhood characteristics they believe impact outcomes. This result has broad implications for the way program effects are used to learn about parameters of other models (Heckman (2010), Rosenzweig and Wolpin (2000)). It is not only the case that researchers should use explicit models, but further that they must do so, even when aided by experimental data.

Using program effects to indirectly draw conclusions about neighborhood effects is shown to be at a clear disadvantage when compared to directly specifying and estimating a model of neighborhood effects. The indirect approach in the literature obfuscates central identifying assumptions necessary for learning about neighborhood effects. By implicitly defining neighborhood quality as a binary variable, this approach has answered the ill-posed question: Do neighborhoods matter for outcomes? Directly specifying and estimating a model of neighborhood effects allows the researcher to relate changes in specific neighborhood characteristics to changes in outcomes. The clear advantage of this approach is that it addresses the well-posed question: How do neighborhoods matter for outcomes?

The paper concludes with a brief discussion of what these considerations imply for the use of evidence from MTO as a test of Wilson (1987). ITT and TOT program effects from MTO should not be interpreted as tests of Wilson (1987). Furthermore, researchers must carefully consider the dynamic nature of Wilson's hypothesis before determining how specific neighborhood effects identified by MTO can be viewed as evidence on Wilson (1987).

The paper proceeds as follows: Section 2 describes the MTO experiment. Section 3.1 states alternative assumptions allowing for the definition of treatment effect parameters in a joint model of potential outcomes and selection into treatment. Section 3.2 discusses how assumptions placed on what we do not observe change the interpretation of these parameters, and Section 4 specifies the additional assumptions used to identify parameters. These Sections draw heavily from results in Heckman and Vytlacil (2005) and Heckman et al. (2006). Section 5 discusses identifying assumptions when treatment is defined as moving through the MTO program. Section 6 considers identifying assumptions when treatment is defined as moving to a high quality neighborhood, presenting empirical evidence on the neighborhood mobility induced by MTO. Section 7.1 discusses the way program effect parameter estimates have been used to indirectly draw conclusions about neighborhood effects, and Section 7.2 discusses implications for using the results from MTO to test Wilson (1987). Section 8 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 gave families awarded Section 8 public housing vouchers the ability to use them beyond the territory of CHA, giving families the option 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)).

The initial relocation process of the Gautreaux program created a quasi-experiment, and its results indicated housing mobility could be an effective policy. 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 city movers (Rosenbaum (1995)).[1]

MTO was designed to replicate these beneficial effects, offering housing vouchers to eligible households between September 1994 and July 1998 in 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)).

Families were drawn from the MTO waiting list through a random lottery. After being drawn, families were randomly allocated into one of three treatment groups. 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.[2]

3 The Definition of Causal Effects

3.1 A Joint Model of Potential Outcomes and Selection

In order to think about effects from MTO, we now define several treatment effect parameters within a standard model of potential outcomes (Rubin (1974), Holland (1986), Heckman and Vytlacil (2005)). Let $Y(1)$ and $Y(0)$ be random variables associated with the potential outcomes in the treated and untreated states, respectively, at the individual level. D is a random variable indicating receipt of a binary treatment, where

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

The measured outcome variable Y is

$$Y = DY(1) + (1 - D)Y(0) \tag{2}$$

[1] 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)).

[2] Section 8 vouchers pay part of a tenant's private market rent. Project-based assistance gives the option of a reduced-rent unit tied to a specific structure.

where potential outcomes are a function of observable characteristics X_D and some treatment level specific unobservable component U_j for $j \in \{0, 1\}$:

$$\begin{aligned} Y(0) &= \mu_0(X_0) + U_0 \\ Y(1) &= \mu_1(X_1) + U_1. \end{aligned} \tag{3}$$

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:

$$Z = \begin{cases} 1 & \text{if treatment is assigned;} \\ 0 & \text{if treatment is not assigned.} \end{cases} \tag{4}$$

Noting it need not be true that $D = Z$, we write $D(Z)$ to denote the treatment received when assigned treatment Z and we explicitly model how individuals select into treatment. We suppose there is a latent index D^* that depends on observable characteristics X, assigned treatment Z, and some unobserved component V as follows:

$$\begin{aligned} D^* &= \mu_D(X_0, Z) - V \\ &= \mu_X(X_0) + \gamma Z - V, \end{aligned} \tag{5}$$

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

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

Finally, define the propensity score conditional on Z to be $\pi^Z(X) \equiv F_V(\mu_D(X, Z)) \equiv Pr(D = 1|X, Z)$.

We follow Heckman and Vytlacil (2005) and Heckman et al. (2006) and assume:

A1 $\gamma_i = \gamma$ for all i and $\gamma \neq 0$

A2 $\{U_0, U_1, V\} \mid X \perp\!\!\!\perp Z$

A3 The distribution of V is continuous

A4 $E[Y(0) \mid X] < \infty$ and $E[Y(1) \mid X] < \infty$

A5 $0 < Pr(D = 1|X) < 1$ for all X

A6 $X = X_1 = X_0$ almost everywhere

Given this joint model of potential outcomes and selection into treatment, there are several treatment effect parameters we might be interested in investigating. We define Intent-to-Treat

(ITT), Treatment-on-the-Treated (TOT), and Local Average Treatment Effect (LATE) parameters:

$$\triangle^{ITT}(x, \pi^0(x), \pi^1(x)) \equiv E[Y \,|\, x,\, Z = 1] - E[Y \,|\, x,\, Z = 0] \tag{7}$$

$$\triangle^{TOT}(x) \equiv E[Y(1) - Y(0) \mid x,\, D = 1] \tag{8}$$

$$\triangle^{LATE}(x, \pi^0(x), \pi^1(x)) \equiv E[Y(1) - Y(0) \mid x,\, D(1) - D(0) = 1] \tag{9}$$

Heckman and Vytlacil (2005) show that these and all of the remaining treatment effect parameters in the literature can be written as weighted averages of a parameter introduced by Björklund and Moffitt (1987), the Marginal Treatment Effect (MTE), which is defined as:

$$\triangle^{MTE}(x, v) \equiv E[Y(1) - Y(0) \,|\, x, v]. \tag{10}$$

We also define $U_D = F_{V|X}(V|X)$, so we can refer interchangeably to $\triangle^{MTE}(x, u_D)$, the MTE at the conditional quantiles of V. The parameters defined in 7 and 9 can be written as averaged MTEs as follows:

$$\triangle^{ITT}(x, \pi^0(x), \pi^1(x)) = \int_{\pi^0(x)}^{\pi^1(x)} \triangle^{MTE}(x, u_D) du_D \tag{11}$$

$$\triangle^{LATE}(x, \pi^0(x), \pi^1(x)) = \frac{1}{\pi^1(x) - \pi^0(x)} \int_{\pi^0(x)}^{\pi^1(x)} \triangle^{MTE}(x, u_D) du_D. \tag{12}$$

Equations 11 and 12 allow us to see the LATE parameters as the average MTE for different combinations of the groups of compliers, always-takers, never-takers, and defiers.[3] Specifically, given a monotonicity assumption to be discussed later, the LATE parameter is the average MTE for compliers.

3.2 Assumptions about the Distribution of Unobservables

Note that so far we have stated no assumption on the relationship between the unobservable components determining potential outcomes and selection into treatment. The treatment effects we have defined in Equations 7-9 exist regardless of the relationship between potential outcomes and V. However, the interpretation of the treatment effect parameters will be very different depending on the assumptions we make about the relationship between the unobservables in the model.

3.2.1 Assumptions within Individuals

Strong ignorability is a standard assumption made in the statistics and econometrics literature about the relationship between the unobservable component determining selection into treatment and those determining potential outcomes. Strong ignorability is fundamentally an assumption about what the econometrician is able to observe; it is that the econometrician can observe all characteristics connecting selection into treatment with treatment effect heterogeneity. Although

[3] See Table 1 or Angrist et al. (1996) for the definition of these groups.

this assumption may be unrealistic in many applications, it is adopted frequently because it is helpful for identification for reasons that will be discussed shortly.

An implication of strong ignorability is that conditional on observables, selection into treatment is not related to treatment effect heterogeneity. Formally, strong ignorability can be written in our model as

SI $\{U_1, U_0\} \perp\!\!\!\perp V \mid X$.

Under SI the MTE is the same for all V. Since the MTE is homogeneous,

$$\triangle^{MTE}(x, u_D) = \triangle^{TOT}(x) = \triangle^{LATE}(x, \cdot, \cdot) \tag{13}$$

for all $u_D \in [0, 1]$ and for all x in the support of X.

Imbens and Angrist (1994) showed it is possible to identify an interpretable parameter, the LATE, even if strong ignorability fails. Recent work in Heckman and Vytlacil (2005), Heckman et al. (2006), and Carneiro et al. (2011) has further defined and estimated treatment effect parameters when relaxing the assumption of strong ignorability by assuming that unobservable treatment effect heterogeneity is related to the unobservable determinants of selection into treatment. Formally, the assumption of essential heterogeneity is that

EH $COV(U_1 - U_0, V) \mid X \neq 0$.

Figure 1 helps to illustrate the implications of SI and EH. The top panel in the figure shows that average treatment effects are allowed to vary across observable characteristics. SI and EH characterize different scenarios once we select a particular value of observable characteristics, x^*. In the middle panel of the figure we see a scenario of SI. The distributions of the potential outcomes must be independent of V given x^*, so the levels of the potential outcomes must be constant across V given x^*. The differences between these levels, the MTEs, are thus constant for all V given x^*.

The bottom panel of Figure 1 shows a contrasting scenario of EH. In this scenario the difference $U_1 - U_0$ is correlated with V, resulting in MTEs that vary across V. In the example displayed the effect of treatment is large for low levels of V, while for large values of V the effect of treatment decreases. Given our latent index model, this implies that for the given observable characteristics x^*, treatment effects are large for those who would be most likely to select into the program and small for those who are more difficult to induce into the program. Finally, Figure 2 shows that while SI and EH are mutually exclusive, they are not exhaustive since individuals might select on the level while not selecting on the gain.

The contrast in the role of instrumental variables under SI versus EH is shown clearly in Figure 1. Under SI it does not matter who is induced into treatment by the instrument since all variation from Z identifies the same homogeneous parameter. Unlike EH, one might assume SI and estimate parameters without the existence of an instrument, perhaps implemented with propensity score matching. In fact, it may appear to be superfluous to use an instrument in conjunction with the SI

assumption. This is not necessarily the case, though, as adding a valid instrument Z to the latent index in Equation 5 can make SI a more plausible assumption.

In contrast to SI, under EH the selection into treatment induced by the instrument is of central interest for interpreting parameters. Since MTEs vary over the support of U_D, the subinterval induced into treatment by the instrument will determine the parameter(s) identified by the instrument. Different instruments that induce different intervals of U_D into treatment will identify different parameters.

3.2.2 Assumptions across Individuals

The parameters in Section 3.1 are all defined conditional on the joint distribution (U, V) where we define $U \equiv (U_0, U_1)$. SI and EH are assumptions about these random variables within individuals, but not across individuals. Assumptions about how these random variables interact across individuals have implications for the joint distribution (U, V) and will change the interpretation of the parameters we have defined.

One possibility satisfying A6 is for X to be a bundle of individual level characteristics including baseline neighborhood characteristics, with one element captured in the unobservables V being peer effects on the selection decision.[4] We now take some terminology from Sobel (2006) to consider the implications of changes to the distribution of V. We suppose the MTO experiment involves N individuals, that there are k_1 people assigned to $Z = 1$, and that $k_0 = N - k_1$ are assigned to $Z = 0$. Let $R(k_0, k_1)$ denote the set of possible realizations of such a randomization, with $r \in R(k_0, k_1)$ denoting one possible realization. If peer effects determining selection into treatment are a part of V, then different realizations r may result in different distributions of V, which we write as $F_{V|r}$. Returning to the fact that all of the parameters defined in Section 3.1 are defined assuming some distribution of (U, V), this implies that these parameters might be very different for some realization r compared to another realization r' (Sobel (2006)).

A standard assumption on the nature of peer effects resolves this problem by ensuring the effects defined in Section 3.1 are the same for all realized random assignments r. This assumption simply assumes there are no peer effects at all. In the context of our model, Angrist and Imbens (1995) state the Stable Unit Treatment Value Assumption (SUTVA) from Rubin (1978) as

SUTVA (a) $V_i \perp\!\!\!\perp Z_j$ for all $j \neq i$

SUTVA (b) $(U_{0i}, U_{1i}) \perp\!\!\!\perp Z_j$ and $(U_{0i}, U_{1i}) \perp\!\!\!\perp D_j$ for all $j \neq i$

Note that SUTVA is an assumption across different individuals, while A2 is an assumption within individuals.

A less restrictive assumption on peer effects that still keeps the effects in Section 3.1 identical across realizations of the randomization is that the distribution of peer effects will be identical under all realizations r. We label this as the Stable Peer Effects Assumption (SPEA):

[4] See page 677 of Heckman and Vytlacil (2005) for a relevant discussion of A6, and see Brock and Durlauf (2007) for a related model of peer effects on the selection decision.

SPEA $(U, V) \perp\!\!\!\perp R$

Note that neither SPEA nor SUTVA is necessary to estimate the parameters defined in Section 3.1, but the model illustrates how the lack of either assumption dramatically changes their interpretation. Since the distribution of peer effects included in V might change in different contexts, this could have very important consequences (Sobel (2006)). We will assume SPEA for the remainder of the analysis, and understanding the types of social interaction allowed under SPEA appears to be a subject for future research.

4 The Identification of Causal Effects

Given the model discussed in Section 3.1 we would ideally be able to identify all MTEs in the support of X and V under assumption EH. In the case that all of the identified MTEs were constant in V conditional on X, we could then proceed under the more restrictive assumption SI. Since data requirements will typically determine both the parameters that we are able to estimate and the assumptions under which we can estimate those parameters, we now consider the parameters identified under EH given various data constraints.

4.1 MTEs

In a more general case than MTO, Z is one or more continuous instruments, allowing us to define $\pi(X, Z) \equiv Pr(D = 1|X, Z)$ and to redefine the parameters in 7-9 by replacing $\pi^Z(X)$ with $\pi(X, Z)$. In such a model Heckman and Vytlacil (1999) develop the method of Local Instrumental Variables (LIV), which is built around the result that

$$\triangle^{MTE}(x, u_D = p) = \frac{\partial E[Y \mid X = x, \pi(X, Z) = p]}{\partial p}. \tag{14}$$

Together with the right hand side of 14, the variation in $\pi(X, Z)$ induced by the continuous instruments can be used to identify $\triangle^{MTE}(x, p)$ for all p in the empirical support of $\pi(X, Z)$. Using this method under both parametric and semiparametric estimation techniques, Carneiro et al. (2011) find that the MTE of attending college on wages is decreasing in U_D for a sample of white males.

Heckman et al. (2006) develop the case of a multi-valued treatment, where $D = j \in \{1, \ldots, J\}$. In their theoretical analysis, Heckman et al. (2006) show that the method of LIV can be extended to the multi-valued case if there exists a set of $J - 1$ instruments each of which exogenously varies one margin of choice while leaving all other margins of choice unaffected. Building on Heckman et al. (2006), Aliprantis and Richter (2012) develop a feasible alternative to LIV when $J - 1$ level-specific instruments are not available. The key insight in Aliprantis and Richter (2012) is that u_D can be identified in an ordered choice model. This allows for the identification of $E[Y(j)|x, u_D]$ for some area in the support of $X \times [0, 1]$ for each j, from which MTE estimates can be constructed.

4.2 Average MTEs

In the case of both MTO and the model we have considered to this point there is a binary instrument.[5] Such a binary instrument is conducive to the estimation of the average MTE over some interval that is determined by selection into treatment. These average MTEs are the parameters defined in Equations 7-9, and they will be identified using some version of the Wald estimator:

$$\frac{E[Y|x, Z=1] - E[Y|x, Z=0]}{E[D|x, Z=1] - E[D|x, Z=0]}.$$

We begin by noting that by comparing mean outcomes at two different values of the instrument we can identify the $\triangle^{ITT}$ parameter simply by assuming A4, which ensures the parameter is finite:

$$\begin{aligned}\triangle^{ITT}(x, \pi^0(x), \pi^1(x)) &\equiv E[Y \mid x,\, Z=1] - E[Y \mid x,\, Z=0] \\ &= E[D(1)Y(1) + (1-D(1))Y(0) \mid x,\, Z=1] \\ &\quad - E[D(0)Y(1) + (1-D(0))Y(0) \mid x,\, Z=0].\end{aligned}$$

If we are further willing to assume A2, then comparing mean outcomes at two different values of the instrument yields a weighted average of the effect on those who select into the program and the effect on those who select out of the program:

$$\begin{aligned}\triangle^{ITT}(x, \pi^0(x), \pi^1(x)) &= E[D(1)Y(1) + (1-D(1))Y(0) \mid x,\, Z=1] \\ &\quad - E[D(0)Y(1) + (1-D(0))Y(0) \mid x,\, Z=0] \\ &= E[(D(1)-D(0))(Y(1)-Y(0)) \mid x] \qquad (15) \\ &= Pr[D(1)-D(0)=1 \mid x]\ E[Y(1)-Y(0) \mid x,\, D(1)-D(0)=1] \qquad (16) \\ &\quad + Pr[D(1)-D(0)=-1 \mid x]\ E[Y(0)-Y(1) \mid x,\, D(1)-D(0)=-1].\end{aligned}$$

The restrictions our assumptions place on the selection model ensure we can identify parameters of interest from Equation 16.[6] Assumption A1 is a monotonicity assumption, ruling out cases in which similar manipulations of the instrument cause some individuals to select into treatment while causing others to select out of treatment. Thus we can assume without loss of generality that $\gamma > 0$, so $Pr[D(1)-D(0)=-1 \mid x] = 0$ and $Pr[D(1)-D(0)=1 \mid x] \neq 0$. Since $D \in \{0,1\}$,

$$\begin{aligned}Pr[D(1)-D(0)=1|x] &= Pr[D(1)=1|x] - Pr[D(0)=1|x] \\ &= E[D|x, Z=1] - E[D|x, Z=0]. \qquad (17)\end{aligned}$$

Substituting 17 into Equation 16, A1 implies we can identify $\triangle^{LATE}(x, \pi^0(x), \pi^1(x))$ by comparing

[5] The MTO instrument technically has three levels, but we abstract from this for the sake of exposition.

[6] Vytlacil (2002) and Vytlacil (2006) show that the identifying assumptions in models with essential heterogeneity are equivalent to the original identifying assumptions for LATEs and generalized LATEs as presented in Imbens and Angrist (1994) and Angrist and Imbens (1995).

those in the data with different values of Z:

$$\frac{E[Y|x, Z=1] - E[Y|x, Z=0]}{E[D|x, Z=1] - E[D|x, Z=0]} = E[Y(1) - Y(0) \mid x, D(1) - D(0) = 1] \tag{18}$$
$$\equiv \triangle^{LATE}(x, \pi^0(x), \pi^1(x)).$$

An additional restriction we might place on the choice model could be

A5* $Pr[D(1) = 1|x] > 0$ and $Pr[D(0) = 1|x] = 0$.

Under A5*

$$D(1) - D(0) = 1|x \iff D(1) = 1|x \ , \tag{19}$$

and we can use 19 to rewrite Equation 16 as

$$\frac{E[Y|x, Z=1] - E[Y|x, Z=0]}{E[D|x, Z=1] - E[D|x, Z=0]} = E[Y(1) - Y(0) \mid x, D = 1] \tag{20}$$
$$\equiv \triangle^{TOT}(x) = \triangle^{LATE}(x, 0, \pi^1(x)).$$

Since Z was randomly allocated in MTO, one option for estimating the unconditional LATE is to simply estimate a TSLS regression without covariates. Frölich (2007) discusses both nonparametric and parametric methods for estimating conditional LATEs.

5 What Program Effects Are Identified by MTO?

Since the model defined in Section 3.1 is built around selection into treatment, it is not fully specified without first defining treatment. Unobservables will be different for different definitions of treatment, and thus our assumptions will change based on our definition of treatment. We now consider identifying assumptions under two definitions of treatment that correspond to effects we hope the MTO experiment will help us to understand.

One obvious definition of treatment we might wish to consider is:

D1 Treatment is moving with the aid of the program (ie, using an MTO voucher).

Under A4 we can identify the ITT parameter by comparing the expected value of the outcome for those assigned to different voucher groups:

$$E[Y \mid x,\ Z = 1] - E[Y \mid x,\ Z = 0] = \triangle^{ITT}(x, \pi^0(x), \pi^1(x)).$$

Under either assumptions (A1-A6, SI, D1) or assumptions (A1-A6, A5*, SI, D1) the Wald estimator allows us to identify the homogeneous program effect of MTO:

$$\frac{E[Y|x, Z=1] - E[Y|x, Z=0]}{E[D|x, Z=1] - E[D|x, Z=0]} = \triangle^{MTE}(x, \cdot) = \triangle^{TOT}(x) = \triangle^{LATE}(x, \cdot, \cdot) \tag{21}$$

If we relax SI by assuming EH, then under (A1-A6, EH, D1) MTO identifies the following program effect that is determined in part by selection into treatment:

$$\frac{E[Y|x, Z=1] - E[Y|x, Z=0]}{E[D|x, Z=1] - E[D|x, Z=0]} = \triangle^{LATE}(x, \pi^0(x), \pi^1(x)). \tag{22}$$

And under (A1-A6, A5*, EH, D1) MTO identifies the following program effect that is also dependent on selection into treatment:

$$\frac{E[Y|x, Z=1] - E[Y|x, Z=0]}{E[D|x, Z=1] - E[D|x, Z=0]} = \triangle^{TOT}(x) = \triangle^{LATE}(x, 0, \pi^1(x)). \tag{23}$$

Since assumptions (A1-A6, A5*, EH, D1) appear reasonable together, the program effect in Equation 23 is identified by MTO. However, this parameter will not be experiment invariant unless an assumption also holds that restricts the permissible types of peer effects. It is unclear whether it is also appropriate to adopt an assumption such as SPEA or SUTVA, and those interested in this issue are directed to the careful discussions in Sobel (2006) and Ludwig et al. (2008).

Estimates of these program effects can be found in the literature on MTO. Some of the major findings are that 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 were, however, positive program 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)). Sanbonmatsu et al. (2006) find program effects on reading scores, math scores, behavior problems, and school engagement that are statistically indistinguishable from zero for MTO children who were 6-20 on December 31, 2001. And perhaps the most surprising result was that while the program improved outcomes for young females, MTO had negative TOT effects on the outcomes of young males (Kling et al. (2007a), Kling et al. (2005)).

6 What Neighborhood Effects Are Identified by MTO?

Another treatment whose effects we might be interested in understanding is defined as follows:

D2 Treatment is moving to a high-quality neighborhood.

Note that under alternative definitions of treatment the selection model in Equations 5 and 6 will be modeling fundamentally different choices. The choice in the selection model under D2 is whether to move to a neighborhood with particular characteristics, while under D1 the choice modeled is whether to move with an MTO voucher. The corresponding change in effect parameters in the model is to effects from moving to neighborhoods of varying quality. In the literature evidence pertaining to parameters of the model under D1 has been presented in discussions on parameters under D2, and vice-versa, showing the importance of clearly stating which modeling assumptions are being made.

6.1 Defining Neighborhood Quality and Assumption A2

6.1.1 Dichotomizing a Continuous Treatment

There are two key reasons unobservables might be correlated with the instrument, which violates assumption A2, and both reasons are related to how we choose to define neighborhood quality in D2. The first problem results from assuming neighborhood quality is a binary variable when it is in fact multi-valued or continuous. For the sake of implementation we might assume

NQB Neighborhood quality D is a binary function of a latent index of neighborhood quality q: $D = \mathbf{1}\{q \geq q^*\}$

To see the problems resulting from dichotomizing neighborhood quality when it is truly multi-valued or continuous, consider an example in which treatment is defined as moving to a neighborhood at the 80th percentile of neighborhood quality or higher (ie, $q^* = 80$). A household that would move to a neighborhood with quality at the 82nd percentile 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 of higher quality, say at the 90th percentile, after being assigned treatment. If this instrument-induced move were to impact outcomes, then U_0 would be correlated with Z. Such a violation of A2 results from the fact that changes in treatment intensity across margins other than those defining the binary treatment affect outcomes.[7]

One way to resolve this issue is to generalize the model in Section 3.1 along the lines developed in Heckman et al. (2006). In the generalized framework we would assume

NQJ Neighborhood quality D is a multi-valued function of a latent index of neighborhood quality q: $D = j \times \mathbf{1}\{C_{j-1} < q \leq C_j\}$ where $j \in \{1, \ldots, J\}$

Given J levels of treatment, there should be some J large enough so that a generalized version of A2 holds.

6.1.2 Projecting Multiple Variables onto a Single Dimension

The second reason unobservables might be correlated with the instrument arises if neighborhood quality is assumed to be represented by one vector when it is in fact multivariate. In the models currently estimated in the literature this assumption is operationalized as:

NQP Neighborhood quality q is a one-dimensional vector that is a scalar function of neighborhood poverty p: $q = \alpha p$

For example, Kling et al. (2007a) estimate neighborhood effects from MTO using a model assuming D2, NQJ, and NQP where MTEs are constant across unobservables.[8]

[7] A discussion related to Assumption NQB can also be found in Angrist and Imbens (1995).

[8] To be precise, the model in Kling et al. (2007a) is the limit of this model as $J \to \infty$. Ludwig and Kling (2007) estimate a similar model with poverty replaced by beat crime rate. MTEs in these analyses are constant in U under the specification in Equation 3 since they assume $U_j = U$ for all $j \in \{1, \ldots, J\}$, so $U_{j+1,i} - U_{j,i} = U_i - U_i = 0$.

If neighborhood quality is truly multivariate, then there might be some neighborhood characteristics affecting outcomes other than poverty. If these characteristics are not perfectly correlated with poverty, then the U_j might be correlated with the instrument Z. Consider an example in which the neighborhood unemployment rate impacts labor market outcomes, with $D \in \{1, \ldots, 10\}$, and $D = j$ if the poverty rate is in the interval $[100 - 10j, 100 - 10(j-1)]$. There is some distribution of unemployment rates for those living in high ($D = j - 1$) and low poverty ($D = j$) neighborhoods, (U_{j-1}, U_j). If the people induced to move into low poverty neighborhoods due to the instrument tend to move to neighborhoods with higher unemployment rates than those who move to low poverty neighborhoods without the instrument, then the distribution of U_j will be different for those with $Z = 0$ than for those with $Z = 1$.

Assumption NQP rules out this possibility. If poverty were perfectly correlated with the unemployment rate, then in this example moving to a low poverty neighborhood would imply moving to a neighborhood with a given unemployment rate regardless of the instrument value, ensuring the distribution of the U_j would not be correlated with Z. Empirical evidence related to NQP is presented in Section 6.2.

A generalization of NQP is:

NQK Neighborhood quality q is a one-dimensional vector that is a linear combination of K observable neighborhood characteristics: $q = \alpha_1 X_1 + \cdots + \alpha_K X_K$

Assumption A2 might be more plausible under NQK than NQP since it uses more information about a neighborhood to determine its quality than solely its poverty rate.

6.2 Empirical Evidence on Assumptions A5, NQP, and NQK

6.2.1 Data

The first source of data used to examine the stated identifying assumptions is the MTO Interim Evaluation sample. The sample contains variables listing the census tracts in which households lived at both the baseline and in 2002, the time the interim evaluation was conducted. These census tracts are used to merge the MTO sample with decennial census data from the National Historical Geographic Information System (NHGIS, Minnesota Population Center (2004)), which provide measures of neighborhood characteristics. These measures are analyzed both as raw values and as the percentiles of the national NHGIS variables from the 2000 census. The variables created in this way include the poverty rate, the percent of adults who hold a high school diploma or a BA, the male Employed-to-Population Ratio (EPR), the share of households with own-children under the age of 18 that are single-headed, and the female unemployment rate.

This analysis focuses on the adults in the MTO Interim Evaluation sample. Weights are used in constructing all estimates.[9]

[9] Weights are used 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

6.2.2 The Neighborhood Mobility Induced by MTO

Consider the generalized model in which neighborhood quality is defined under assumptions D2, NQJ, and NQK with $j \in \{1, \ldots, 10\}$ and

$$D = j \times \mathbf{1}\{10 \times (j-1) < q \leq 10 \times j\},$$

where q is the percentile of neighborhood quality. A key assumption that can be empirically tested under this definition is A5, which is an assumption about the observed treatment states. The generalized version of assumption A5 is that $0 < Pr(D = j|X) < 1$ for all X, or that there are some persons in each treatment state.

Given the difficulties related to assumption NQP discussed in Section 6.1.2, we adopt NQK by combining several measures of neighborhood quality into a single vector representing neighborhood quality. Principal components analysis is used to determine which single vector combines the most information about the national distribution of the poverty rate, the percent with high school degrees, the percent with BAs, the percent of single-headed households, the male EPR, and the female unemployment rate. Table 2 shows that the resulting univariate index explains 63 percent of the variance of these neighborhood characteristics, and that no additional eigenvector would explain more than 13 percent of the variance of these variables. Table 3 displays the coefficients relating each of these variables to the index vector. Relevant for deciding between assumptions NQP and NQK, the magnitudes of the coefficients for most variables are similar to the magnitude of the coefficient for poverty.

Figure 3a shows the expected negative correlation between neighborhood quality and neighborhood poverty rate. We can see in Figure 3b that the US population distribution of neighborhood poverty rates in 2000 had a long right tail. Similarly, Figure 3c shows that the US population distribution of neighborhood quality had a long left tail in 2000. Figures 3d and 3e show how far in the tails of these national distributions much of the MTO sample typically resided.

Moving from a neighborhood with a poverty rate of 70 percent to a neighborhood with a 50 percent poverty rate might be a large change in the poverty rate, but how big is this change relative to the national distribution of neighborhoods? That is, how much of a change in quality does this 20 percent change represent given a starting rate of 70 percent? An alternative way of measuring poverty and quality that addresses this question is to use the ranking of neighborhoods relative to those of the rest of the US population. These measures are shown for the entire US population in Figure 4a. What we can see is that although the expected negative relationship still remains, there is now considerable variation in one variable conditional on the other. Consider, for example, that there are neighborhoods with the median poverty rate that are extremely low quality, and neighborhoods with the same poverty rate that are extremely high quality. This level of variation may not be surprising given the coefficients reported in Table 3, and supports the adoption of

time trends. Second, sampling weights must be used to account for the sub-sampling procedures used during the interim evaluation data collection.

assumption NQK over NQP.

Figure 4b shows that very few MTO adults were induced into high quality neighborhoods. At the time of the interim evaluation less than 10 percent of the experimental group lived in neighborhoods whose quality was above the median of the national distribution. It is difficult to know for sure, but it appears reasonable to believe that the analogous distributions from Gautreaux would have had more mass in the right tail of the national distribution of neighborhood quality.[10]

The distributions in Figure 4b can be seen as a violation of the generalized version of assumption A5. While technically true for all j without conditioning on X, for the sake of estimation the generalized version of A5 is only likely to hold for $j \in \{1, \ldots, 5\}$ or $j \in \{1, \ldots, 6\}$. By the time of the interim evaluation less than 20 percent of the MTO experimental group lived in neighborhoods above the 30th percentile of the national distribution of quality, and less than 10 percent lived in neighborhoods above the median.

6.3 The Neighborhood Effects Identified by MTO

Effects from moving to high quality neighborhoods are not identified by MTO. Given the evidence in Section 6.2.2, any definition of treatment of the form D2 would have to restrict measures of quality to the lower half of the national distribution of neighborhood quality to satisfy assumption A5.

Once the focus on quality is restricted to accommodate A5, we can see that A5 appears more reasonable than A5*, as it is likely that some households will move to a relatively high quality neighborhood regardless of whether they receive a voucher through MTO or not. Under assumptions (A1-A6, SPEA, EH, D2-NQB) the Wald estimator identifies the LATE:

$$\begin{aligned}\frac{E[Y|x, Z=1] - E[Y|x, Z=0]}{E[D|x, Z=1] - E[D|x, Z=0]} &= \triangle^{LATE}(x, \pi^0(x), \pi^1(x)) \\ &= \frac{1}{\pi^1(x) - \pi^0(x)} \int_{\pi^0(x)}^{\pi^1(x)} \triangle^{MTE}(x, u_D) du_D. \end{aligned} \tag{24}$$

If we believe assumption A2 will fail to hold when treatment is defined under D2-NQB for the reasons discussed in Section 6.1.1, we could alternatively define treatment under D2-NQJ to

[10] DeLuca and Rosenbaum (2003) find that 66 percent of the suburban group and 13 percent of the city group live in the suburbs of Chicago 14 years after original placement through Gautreaux. DeLuca and Rosenbaum (2003) cite limited availability of housing, and not selection to not move through the program, as the reason only 20 percent of eligible applicants moved through Gautreaux. This claim is based on evidence that 95 percent of participating households accepted the first unit offered to them. Furthermore, it is likely that Gautreaux induced larger changes in school quality than MTO (Rubinowitz and Rosenbaum (2000), p 162). Taken together, this evidence is suggestive that Gautreaux induced more households into high quality neighborhoods than MTO.

generate level j specific analogues to 10 and 24:

$$\triangle^{MTE}_{j,j+1}(x, u_D)du_D = E[Y(j+1) - Y(j)|X = x, U_D = u_D]$$
$$\triangle^{LATE}_{j,j+1}(x, \pi^0_j(x), \pi^1_j(x)) = \frac{1}{\pi^1_j(x) - \pi^0_j(x)} \int_{\pi^0_j(x)}^{\pi^1_j(x)} \triangle^{MTE}_{j,j+1}(x, u_D)du_D.$$

Versions of the model have been estimated in Kling et al. (2007a) and Ludwig et al. (2008) under (A1-A6, SPEA, SI, and D2-NQJ-NQP). A dose-response analysis is used in Kling et al. (2007a) to determine if parameters are constant across all j to $j+1$ transitions in $\{1, \ldots, J\}$. Aliprantis and Richter (2012) estimate the model under (A1-A6, SPEA, EH, D2-NQJ-NQK). That analysis makes A2 more plausible by relaxing D2-NQJ-NQP to D2-NQJ-NQK, and allows for the identification and estimation of MTEs that are heterogeneous over unobservables by relaxing SI to EH.[11]

7 Discussion

7.1 Relating Program Effects from MTO to Neighborhood Effects

As defined above, program effects and neighborhood effects are clearly different parameters defined in distinct models (Heckman (2010)). Yet much of the literature on MTO has drawn conclusions about neighborhood effects based on estimates of program effects from MTO. Conclusions in the literature have typically been reached not by equating these two types of parameters, but rather by combining evidence on program effects from MTO together with logical arguments to indirectly draw conclusions about neighborhood effects.[12]

This Section shows that program effects from MTO should not be used to make indirect inference about neighborhood effects. This Section begins by formally stating the logical argument that has been combined with estimates of the program effects from MTO to make conclusions about neighborhood effects. It is shown that the validity of this argument can only be determined if those neighborhood characteristics thought to affect outcomes are explicitly specified. This fact leaves no theoretical justification for using ITT and TOT program effect estimates to make indirect inferences about neighborhood effects. Such an approach offers no advantage over directly estimating neighborhood effects. Furthermore, it is shown that this logical argument through which program effects from MTO have been used to learn about neighborhood effects answers an ill-posed question. For these reasons, trying to learn about neighborhood effects from MTO by focusing on its program effects obfuscates the evidence on neighborhood effects from MTO.

[11] Note that NQK need not be adopted only in conjunction with NQJ. A version of Assumption NQB-NQK is adopted in Sampson et al. (2008) using a similar index of neighborhood quality to that used in this analysis.

[12] This is the author's current interpretation of the literature, most prominently represented by Kling et al. (2007a) and Ludwig et al. (2008). The distinction between program and neighborhood effect parameters has not always been made clearly, and some studies do seem to equate program effects with neighborhood effects, even when using this indirect logic. Early examples where this distinction is unclear are Ludwig et al. (2001) and Kling et al. (2005), and more recent examples include Sanbonmatsu et al. (2012) and Gennetian et al. (2012).

7.1.1 There is No Advantage from Using Program Effects from MTO for Indirect Inference on Neighborhood Effects

Consider the following three propositions, or statements with truth values:

$$P = \text{“Individual } i \text{ receives an MTO voucher.”}$$
$$Q = \text{“Individual } i \text{ lives in a ‘good’ neighborhood.”}$$
$$R = \text{“Individual } i \text{ is employed.”}$$

It is important to note that “good” neighborhoods and employment both take binary definitions, allowing for the straightforward assignment of truth values to P, Q, and R. We assume P precedes Q and Q precedes R, and we seek to make inference about the causal chain $P \rightarrow Q \rightarrow R$.

We denote the negation of event A to be $\sim A$. We follow Lewis (2001) and define the counterfactual conditional operator $\Box\!\!\rightarrow$ as:

“If it were the case that ____ , then it would be the case that ____ .”

We define causality when event A temporally precedes event B, and when A and $\sim A$ represent counterfactual worlds identical to our own world in all ways except for variable A being manipulated.[13] In this setting we say that for individual i: (1) A causes B if $A \Box\!\!\rightarrow B$ is true and $\sim A \Box\!\!\rightarrow B$ is false, (2) A has no causal effect on B when the truth values of $A \Box\!\!\rightarrow B$ and $\sim A \Box\!\!\rightarrow B$ coincide, and (3) $\sim A$ causes B if $A \Box\!\!\rightarrow B$ is false and $\sim A \Box\!\!\rightarrow B$ is true.[14]

Table 4 presents a truth table illustrating this definition of $\Box\!\!\rightarrow$, and Table 5 presents a truth table illustrating this definition of causal effect. Note that the truth value of each Row in Table 4 pertains to one particular world, but the truth value of each Row in Table 5 pertains to multiple, or counterfactual, worlds. In the following analysis we interpret the truth value of each cell (Row and Column combination) in Tables 5-8 to pertain to the counterfactual world identical to our own world in all ways except that the antecedent has been manipulated to be true while leaving every other variable undisturbed.

Under this definition of causation, the combined truth values of the conditional statements $P \Box\!\!\rightarrow Q$ (“If it were the case that individual i receives an MTO voucher, then it would be the case that individual i lives in a ‘good’ neighborhood.”) and $\sim P \Box\!\!\rightarrow Q$ represent a statement about the causal effects MTO had on residential mobility. The combined truth values of $Q \Box\!\!\rightarrow R$ and $\sim Q \Box\!\!\rightarrow R$ can be interpreted as a statement about neighborhood effects, and the truth values of $P \Box\!\!\rightarrow R$ and $\sim P \Box\!\!\rightarrow R$ represent a statement about program effects.

Table 6 presents a truth table of the conditional statements relevant to the effects of MTO on

[13] Scheines (1997) refers to manipulation of variable A while leaving every other variable undisturbed as ideal manipulation of A.

[14] Defining causality is a non-trivial task that has yet to be entirely resolved by scientists and philosophers (See Scheines (2004) for a discussion.). We adopt this operational definition because it allows us to formalize the interpretation of results from MTO.

residential mobility, neighborhood effects, and program effects on outcomes. Due to the assumed temporal ordering of events and the experimental assignment of MTO vouchers, the truth values in Columns 5 and 6 can be inferred from those in Columns 1-4. Note that the truth table is defined for one individual i, but due to the fundamental problem of evaluation we can only make statements about this truth table for individuals in our sample on average.

We are interested in determining what parts of the truth table correspond to the empirical results from MTO. We can interpret the empirical evidence to rule out Column 1 and Column 2 both being true at the same time, eliminating Rows 1-4 from our consideration. And since we can also interpret the empirical evidence to eliminate Rows 9-12, we can focus our attention on the rows highlighted in the truth table in Table 7.

The most prominent interpretation of the empirical results from MTO in the literature might be summarized as follows: "If neighborhood environments affect behavior ... then these neighborhood effects ought to be reflected in ITT and TOT impacts on behavior" (Ludwig et al. (2008), pp 181-182). This interpretation can be related to the truth table: Most literally, MTO affecting neighborhood quality and neighborhood quality affecting employment is represented by Row 6 or Row 7 of the truth table. If this were the case in reality (ie, reality were represented by Row 6 or Row 7), we would observe program effects from MTO. Less literally, this interpretation of the results from MTO is that MTO's effects on neighborhood quality and the program effect estimates are most accurately represented by Row 8 of the truth table. Therefore, although we do not directly observe neighborhood effects, we have to infer they do not exist.

The justification for using ITT and TOT program effect parameters to make indirect inference about neighborhood effects is that doing so allows the researcher to impose no assumptions about the nature of neighborhood effects. For example, Ludwig et al. (2008) argue that "Randomization eliminates the need to correctly specify which neighborhood characteristic matters for each outcome to learn about whether neighborhoods matter" (p. 151). This is not true: ITT and TOT estimates are not informative about neighborhood effects without the researcher first specifying which neighborhood characteristic matters for each outcome. Without first specifying which neighborhood characteristics qualify as "good," it is not possible to determine the truth value of either $P \Box\!\!\rightarrow Q$ or $\sim P \Box\!\!\rightarrow Q$, and therefore it is not possible to determine whether Rows 5-8 or Rows 13-16 of the truth table more accurately correspond to the empirical evidence from MTO.[15] After specifying which neighborhood characteristics qualify as "good," there is no justification for researchers not to estimate the model under definition D2 so as to provide direct evidence on $Q \Box\!\!\rightarrow R$ and $\sim Q \Box\!\!\rightarrow R$.

[15] This is the central debate between Ludwig et al. (2008) and Clampet-Lundquist and Massey (2008); whether $P \Box\!\!\rightarrow Q$ is true or false. If $P \Box\!\!\rightarrow Q$ is true, then we must be on Row 8 of the truth table, allowing us to infer neighborhood effects do not exist. If $P \Box\!\!\rightarrow Q$ is false, then we could either be on Row 14 or 16 of the truth table, which does not allow us to make any inference about neighborhood effects.

7.1.2 The Focus on Program Effects Obfuscates the Evidence on Neighborhood Effects from MTO

Adopting a definition of "good" in Q is identical to specifying the mechanisms through which researchers believe neighborhoods impact outcomes. A huge disadvantage of the indirect approach above is that it has made this specification implicit, rather than explicit. For example, although Ludwig et al. (2008) and Clampet-Lundquist and Massey (2008) disagree about the truth value of $P \Box\!\!\rightarrow Q$, neither group of researchers explicitly defines "good" in Q. One implicit definition of "good" adopted in these studies is "lower poverty than the MTO control group."

Studies using the program effects from MTO to make indirect inferences about neighborhood effects have implicitly defined neighborhood quality in a binary sense. This has not been clear because the logic has not been explicitly stated in the literature. The analysis in this paper has made clear that neighborhood quality should not be defined as a binary variable taking on only one of two values (ie, $\{Q, \sim Q\}$): Section 6.1.1 discusses clear reasons why this assumption, NQB, must be relaxed. Furthermore, the precise way NQB is relaxed matters. Even allowing neighborhood quality to assume multiple levels, the assumption that these levels can be characterized in one dimension is itself a strong assumption. Section 6.1.2 discussed various approaches to creating a univariate index of quality, and Section 6.2.2 examined empirical evidence on the univariate index created by a statistical procedure to be the optimal from one set of such indeces.

Using program effects from MTO to make indirect inference on neighborhood effects has obfuscated these issues. In particular, by implicitly defining neighborhood quality as a binary variable, it has answered the ill-posed question: Do neighborhoods matter for outcomes?

By relating changes in specific neighborhood characteristics to outcomes, directly estimating and specifying a model of neighborhood effects addresses the well-posed question: How do neighborhoods matter for outcomes? Consider a re-casting of Q into J statements:

$$Q_j = \text{"Individual } i \text{ lives in a neighborhood of quality level j,} \quad \text{for j} = 1, \ldots, \text{J."}$$

As we can see from the truth table in Table 8, the counterfactuals under consideration now become more complicated than a comparison of outcomes under two types of neighborhood environments. In order to make inference about neighborhood effects, the researcher needs to make statements about selection (ie, $P \Box\!\!\rightarrow Q_j$ and $\sim P \Box\!\!\rightarrow Q_j$, each for $j = 1, \ldots, J$.).[16] Only then can the researcher relate those specific changes in neighborhood characteristics from MTO to the changes in outcomes from MTO. The most efficient way of relating the changes in neighborhood characteristics induced by MTO to changes in outcomes is to directly specify and estimate a model of neighborhood effects.

7.2 MTO and Wilson (1987)

Can we test Wilson (1987)'s hypothesis using the results from MTO? Perhaps, but several issues have gone under-appreciated in the literature interpreting the results from MTO as a test of Wilson

[16] This is true even for the indirect approach using program effects.

(1987). First, the literature has interpreted program effects as evidence on neighborhood effects. As shown above, ITT and TOT program effects from MTO should not be interpreted as evidence on neighborhood effects. ITT and TOT program effects from MTO are not relevant for testing Wilson (1987).

Second, the researcher must specify which contemporary neighborhood effects represent a test of Wilson (1987), and why they do so. It is important to remember that Wilson (1987)'s hypothesis is a dynamic one. If residential sorting and secular changes in the labor market did in fact lead some neighborhoods to start declining in the 1960s and 1970s (Collins and Margo (2000), Aliprantis and Carroll (2012), Badel (2010)), then one could easily imagine that the negative externalities experienced near the 20th quantile of today's neighborhood quality might compare with those experienced at much lower quantiles during the 1960s. Secular changes in neighborhoods, and the externalities associated with them, make it difficult to judge what contemporary counterfactuals appropriately test Wilson (1987).[17]

Finally, once the researcher specifies which contemporary neighborhood effects represent a test of Wilson (1987), they must then use the data from MTO to identify them. The empirical evidence presented in this analysis shows that the effects from changes in neighborhood quality identified by MTO are likely to be effects of moves from the very worst neighborhoods in the country to slightly less-bad neighborhoods. As analyzed in Aliprantis and Richter (2012), in practice the MTO data only allow for the identification of effects from moves across the first and second deciles of the 2000 national distribution of neighborhood quality.[18]

These issues have not been sufficiently addressed in the literature on MTO to, at this point, interpret the evidence from MTO as a test of Wilson (1987).

8 Conclusion

This paper has reviewed the assumptions necessary to identify various parameters using the variation in neighborhood of residence induced by the Moving to Opportunity (MTO) housing mobility experiment. An index of neighborhood quality was created that reflects a neighborhood's poverty rate as well as several other characteristics. Empirical evidence was presented that MTO did not induce participants into high quality neighborhoods. One key result of the paper was to show that using MTO voucher assignment as an instrument for neighborhood quality does not identify effects from moving to a high quality neighborhood.

This paper has also shown that the literature draws mistaken conclusions from MTO by using its program effects to learn about neighborhood effects. It was shown how the most prominent

[17] Another difficulty is that neighborhood effects of the type relevant for Wilson (1987) may need a particular time horizon to affect outcomes. The analysis in this paper addresses this issue by classifying mobility from MTO in terms of location at the time of the interim evaluation. See Sampson (2008) for a careful discussion of this issue.

[18] Furthermore, this is also true for neighborhood effect parameters defined in models assuming NQP. For example, Quigley and Raphael (2008) note that "The effect of treatment under the MTO program was, on average, to move households in the five MTO metropolitan areas from neighborhoods at roughly the 96th percentile of the neighborhood poverty distribution to neighborhoods at the 88th percentile" (p 3).

interpretation of results from MTO in the literature uses ITT and TOT program effects from MTO to *indirectly* draw conclusions about neighborhood effects. This interpretation was expressed in terms of formal logic, which showed that such an indirect approach offers no advantage for learning about neighborhood effects over directly estimating them. Even using this indirect approach, the researcher cannot escape explicitly specify those neighborhood characteristics they believe impact outcomes.

Using program effects to indirectly draw conclusions about neighborhood effects was shown to be at a clear disadvantage when compared to directly specifying and estimating a model of neighborhood effects. The indirect approach in the literature allows central identifying assumptions necessary for learning about neighborhood effects to be made implicity. By implicitly defining neighborhood quality as a binary variable, this approach has answered the ill-posed question: Do neighborhoods matter for outcomes? Directly specifying and estimating a model of neighborhood effects allows the researcher to relate changes in specific neighborhood characteristics to changes in outcomes. The clear advantage of this approach is that it addresses the well-posed question: How do neighborhoods matter for outcomes? These results have broad implications for the way program effects are used to learn about parameters of other models.

The analysis ended with a discussion on the use of evidence from MTO as a test of Wilson (1987). It was shown that ITT and TOT program effects from MTO should not be interpreted as tests of Wilson (1987), and that researchers must carefully consider the dynamic nature of Wilson's hypothesis before determining how specific neighborhood effects identified by MTO can be viewed as evidence on Wilson (1987).

References

Aliprantis, D. and D. Carroll (2012). Neighborhood dynamics and the distribution of opportunity. *Federal Reserve Bank of Cleveland Working Paper 12-12r*.

Aliprantis, D. and F. G.-C. Richter (2012). Marginal neighborhood effects from Moving to Opportunity. *Federal Reserve Bank of Cleveland Working Paper 12-08r*.

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.

Badel, A. (2010). Understanding permanent black-white earnings inequality. *Federal Reserve Bank of St. Louis Working Paper 2010-047B*.

Björklund, A. and R. Moffitt (1987). The estimation of wage gains and welfare gains in self-selection models. *The Review of Economics and Statistics 69*(1), pp. 42–49.

Brock, W. and S. Durlauf (2007). Identification of binary choice models with social interactions. *Journal of Econometrics 140*(1), 52–75.

Carneiro, P., J. J. Heckman, and E. J. Vytlacil (2011). Estimating marginal returns to education. *American Economic Review 101*(6), 2754–2781.

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.

Collins, W. and R. Margo (2000). Residential segregation and socioeconomic outcomes: When did ghettos go bad? *Economics Letters 69*(2), 239–243.

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.

Frölich, M. (2007). Nonparametric IV estimation of local average treatment effects with covariates. *Journal of Econometrics 139*(1), 35–75.

Gennetian, L. A., M. Sciandra, L. Sanbonmatsu, J. Ludwig, L. F. Katz, G. J. Duncan, J. R. Kling, and R. C. Kessler (2012). The long-term effects of Moving to Opportunity on youth outcomes. *Cityscape 14*(2), 137–167.

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

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 (1999). Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proceedings of the National Academy of Sciences 96*(8), 4730–34.

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

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

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.

Kneebone, E., C. Nadeau, and A. Berube (2011, November). *The Re-Emergence of Concentrated Poverty: Metropolitan Trends in the 2000s*. Washington, DC: The Brookings Institution.

Lewis, D. (2001). *Counterfactuals* (2nd ed.). Wiley-Blackwell.

Ludwig, J., G. J. Duncan, and P. Hirschfield (2001). Urban poverty and juvenile crime: Evidence from a randomized housing-mobility experiment. *The Quarterly Journal of Economics 116*(2), 655–679.

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

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.

Massey, D. and N. Denton (1993). *American Apartheid: Segregation and the Making of the Underclass*. Cambridge: Harvard University Press.

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.

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

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.

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

Quigley, J. M. and S. Raphael (2008). Neighborhoods, economic self-sufficiency, and the mto program. *Brookings-Wharton Papers on Urban Affairs 8*(1), 1–46.

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.

Rosenzweig, M. R. and K. I. Wolpin (2000). Natural "Natural Experiments" in Economics. *Journal of Economic Literature 38*, 827–874.

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

Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *The Annals of Statistics 6*(1), 34–58.

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.

Sampson, R. J., P. Sharkey, and S. W. Raudenbush (2008). Durable effects of concentrated disadvantage on verbal ability among African-American children. *Proceedings of the National Academy of Sciences of the United States of America 105*(3), 845–852.

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.

Sanbonmatsu, L., J. Marvakov, N. A. Potter, F. Yang, E. Adam, W. J. Congdon, G. J. Duncan, L. A. Gennetian, L. F. Katz, J. R. Kling, R. C. Kessler, S. T. Lindau, J. Ludwig, and T. W. McDade (2012). The long-term effects of Moving to Opportunity on adult health and economic self-sufficiency. *Cityscape 14*(2), 109–136.

Scheines, R. (1997). An introduction to causal inference. In V. McKim and S. Turner (Eds.), *Causality in Crisis?* Univ. of Notre Dame Press.

Scheines, R. (2004). Causation. In *New Dictionary of the History of Ideas*. Charles Scribner & Sons.

Sobel, M. E. (2006). What do randomized studies of housing mobility demonstrate?: Causal inference in the face of interference. *Journal of the American Statistical Association 101*(476), 1398–1407.

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.

Vytlacil, E. (2002). Independence, monotonicity, and latent index models: An equivalence result. *Econometrica 70*(1), 331–341.

Vytlacil, E. (2006). Ordered discrete-choice selection models and local average treatment effect assumptions: Equivalence, nonequivalence, and representation results. *The Review of Economics and Statistics 88*(3), 578–581.

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

Figures

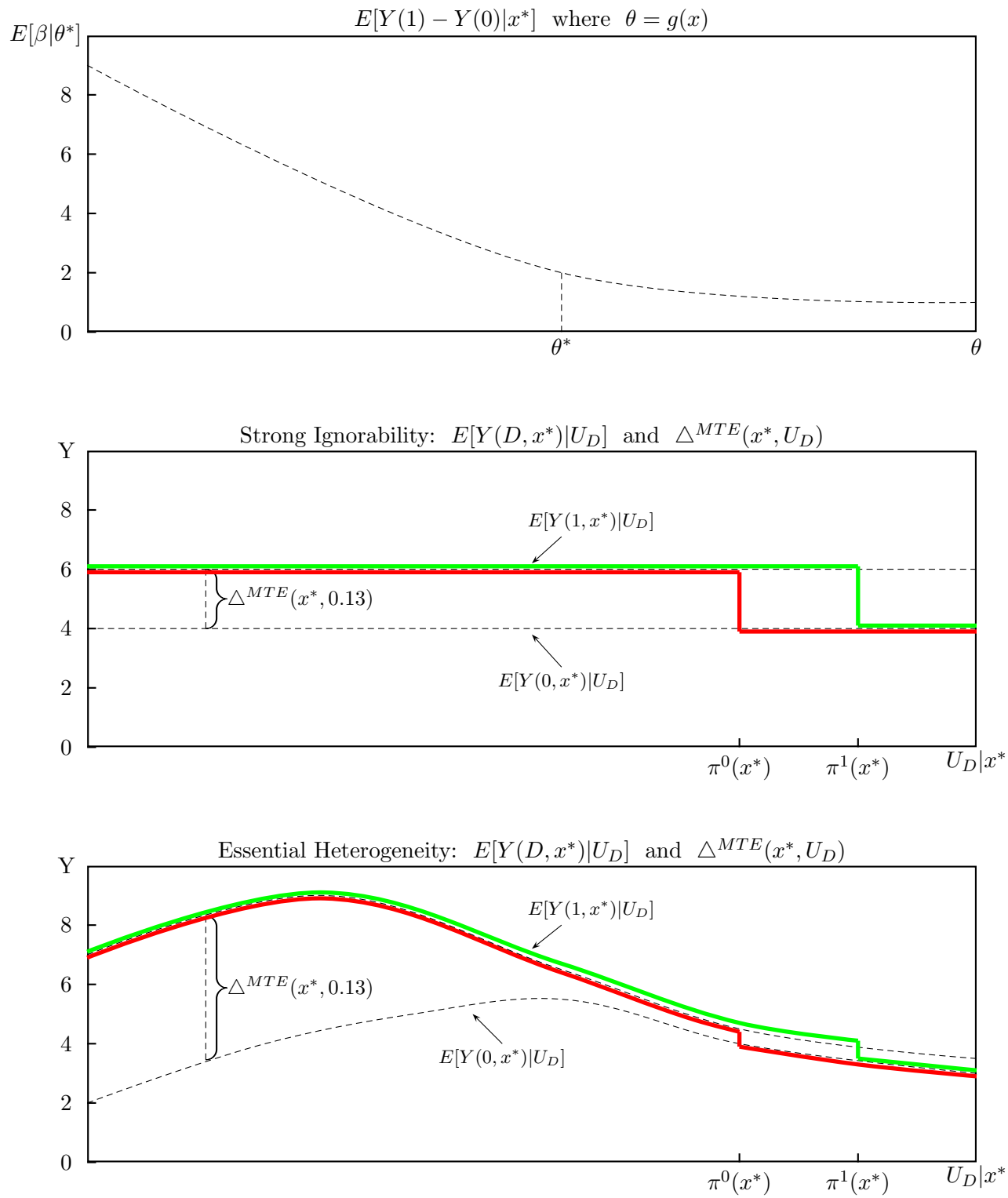


Figure 1: Examples of Strong Ignorability and Essential Heterogeneity

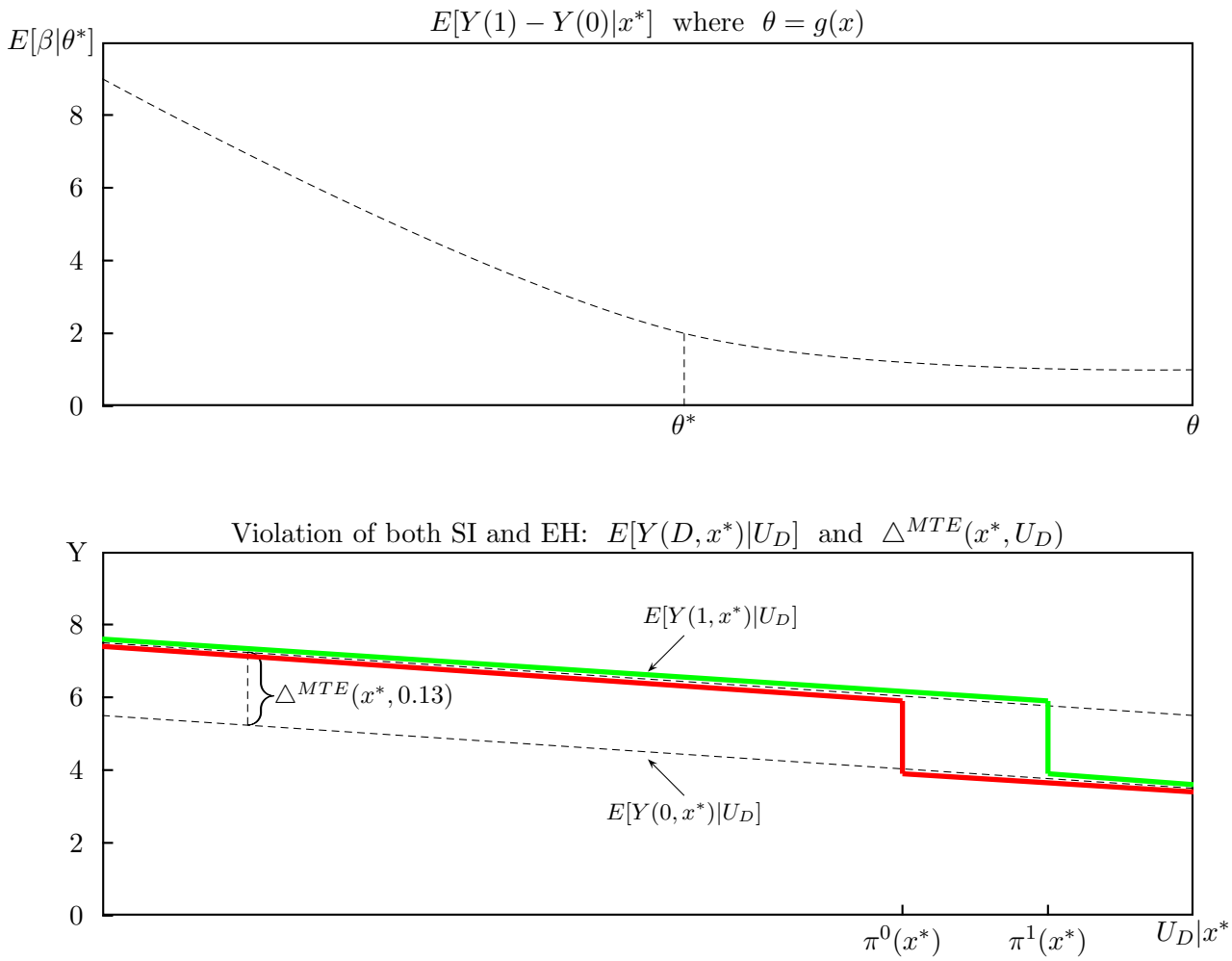
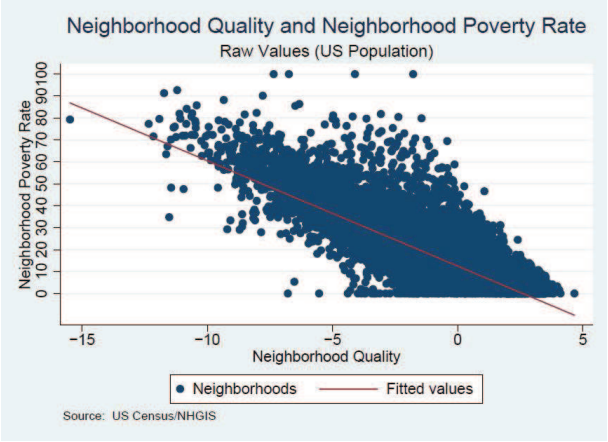
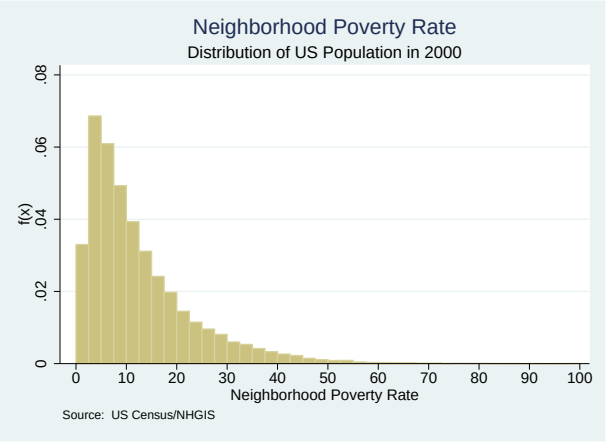


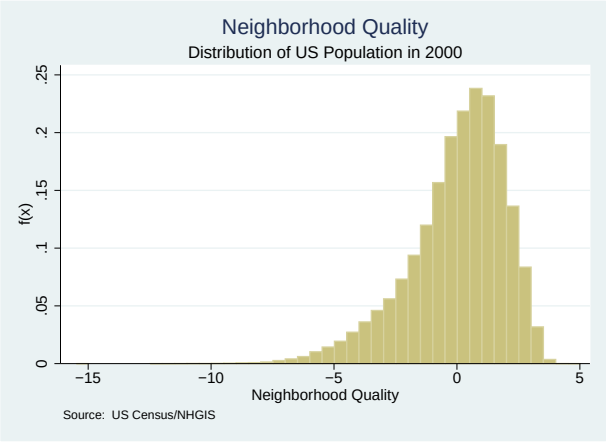
Figure 2: Example Violating Both Strong Ignorability and Essential Heterogeneity



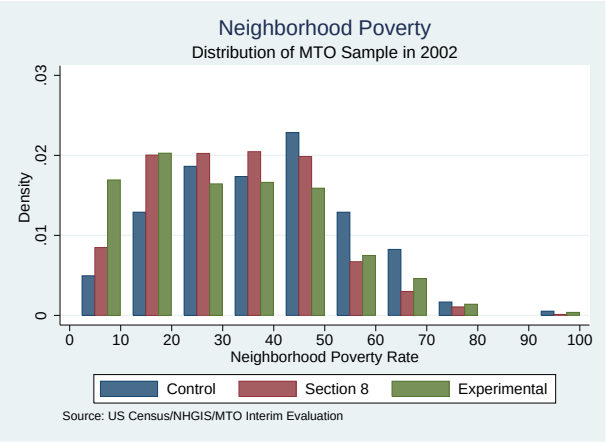
(a) Raw Measures of Neighborhood Quality and Poverty in 2000, US Population



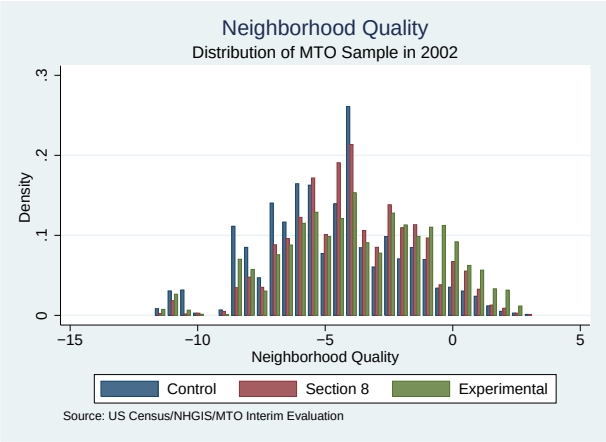
(b) Neighborhood Poverty Rate in 2000, US Population



(c) Raw Measure of Neighborhood Quality in 2000, US Population

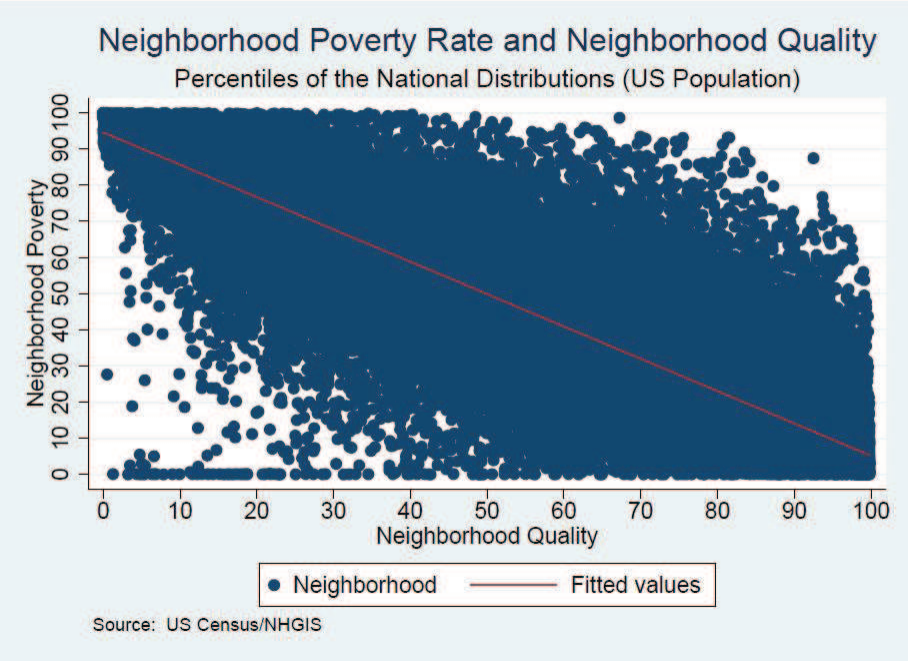


(d) Neighborhood Poverty Rate in 2002, MTO Sample

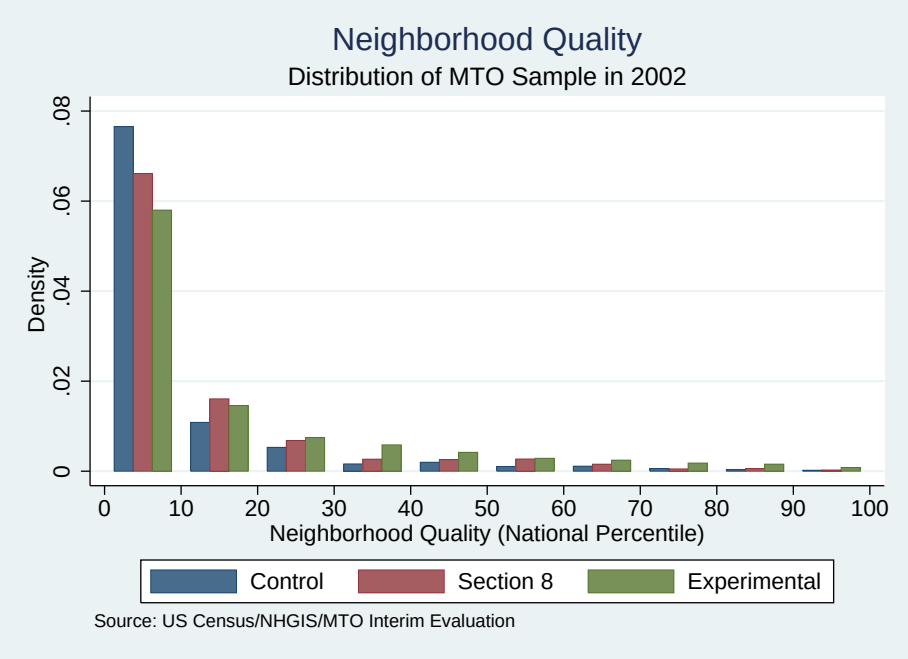


(e) Raw Measure of Neighborhood Quality in 2002, MTO Sample

Figure 3: Neighborhood Poverty Rate and Neighborhood Quality



(a) Percentile Measures of Neighborhood Quality and Poverty in 2000, US Population



(b) Percentile Measure of Neighborhood Quality in 2002, MTO Sample

Figure 4: Neighborhood Poverty and Quality of MTO Participants

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: Proportion of Variance Explained by Principal Components Eigenvectors

Eigenvector	Eigenvalue	Proportion of Variance
1	3.81	0.63
2	0.79	0.13
3	0.56	0.09
4	0.39	0.07
5	0.31	0.05
6	0.14	0.02

Table 3: Principal Components Analysis: First Eigenvector Coefficients

Variable	Coefficient
Poverty Rate	-0.46
HS Graduation Rate	0.44
BA Attainment Rate	0.35
Percent Single-Headed HHs	-0.38
Male EPR	0.41
Female Unemployment Rate	-0.40

Table 4: Truth Table

Definition of Counterfactual Conditional Operator

$\Box\!\!\rightarrow$ reads "If it were the case that ____ , then it would be the case that ____ ."

Column	1	2	3	4	5	6	7
Row	A	B	$\sim A$	$A \rightarrow B$	$\sim A \rightarrow B$	$A \Box\!\!\rightarrow B$	$\sim A \Box\!\!\rightarrow B$
1	T	T	F	T	T	T	·
2	T	F	F	F	T	F	·
3	F	T	T	T	T	·	T
4	F	F	T	T	F	·	F

Table 5: Truth Table

Definition of Counterfactual Conditional Operator

$\Box\!\!\rightarrow$ reads "If it were the case that ____ , then it would be the case that ____ ."

Column	1	2	3	4	5
Row	$A \Box\!\!\rightarrow B$	$\sim A \Box\!\!\rightarrow B$	A causes B	A has no causal effect on B	$\sim A$ causes B
1	T	T	F	T	F
2	T	F	T	F	F
3	F	T	F	F	T
4	F	F	F	T	F

Table 6: Truth Table

Definition of Propositions

P = "Individual i receives an MTO voucher."
Q = "Individual i lives in a 'good' neighborhood."
R = "Individual i is employed."

Definition of Counterfactual Conditional Operator

$\Box\!\!\rightarrow$ reads "If it were the case that ____ , then it would be the case that ____ ."

Column	1	2	3	4	5	6
Row	$P \Box\!\!\rightarrow Q$	$\sim P \Box\!\!\rightarrow Q$	$Q \Box\!\!\rightarrow R$	$\sim Q \Box\!\!\rightarrow R$	$P \Box\!\!\rightarrow R$	$\sim P \Box\!\!\rightarrow R$
1	T	T	T	T	T	T
2			T	F	T	T
3			F	T	F	F
4			F	F	F	F
5	T	F	T	T	T	T
6			T	F	T	F
7			F	T	F	T
8			F	F	F	F
9	F	T	T	T	T	T
10			T	F	F	T
11			F	T	T	F
12			F	F	F	F
13	F	F	T	T	T	T
14			T	F	F	F
15			F	T	T	T
16			F	F	F	F

Table 7: Truth Table

Definition of Propositions

P = "Individual i receives an MTO voucher."
Q = "Individual i lives in a 'good' neighborhood."
R = "Individual i is employed."

Definition of Counterfactual Conditional Operator

$\Box\!\!\rightarrow$ reads "If it were the case that ____ , then it would be the case that ____ ."

Column	1	2	3	4	5	6
Row	$P \Box\!\!\rightarrow Q$	$\sim P \Box\!\!\rightarrow Q$	$Q \Box\!\!\rightarrow R$	$\sim Q \Box\!\!\rightarrow R$	$P \Box\!\!\rightarrow R$	$\sim P \Box\!\!\rightarrow R$
5	T	F	T	T	T	T
6			T	F	T	F
7			F	T	F	T
8			F	F	F	F
13	F	F	T	T	T	T
14			T	F	F	F
15			F	T	T	T
16			F	F	F	F

Table 8: Truth Table

Definition of Propositions

P = "Individual i receives an MTO voucher."
Q_1 = "Individual i lives in a neighborhood of quality 1."

$\vdots$

Q_J = "Individual i lives in a neighborhood of quality J."
R = "Individual i is employed."

Definition of Counterfactual Conditional Operator

$\Box\!\!\rightarrow$ reads "If it were the case that ____ , then it would be the case that ____ ."

Column	1	2	$\cdots$	$2J-1$	$2J$	$2J+1$	$\cdots$	$3J$	$3J+1$	$3J+2$
Row	$P\Box\!\!\rightarrow Q_1$	$\sim P\Box\!\!\rightarrow Q_1$	$\cdots$	$P\Box\!\!\rightarrow Q_J$	$\sim P\Box\!\!\rightarrow Q_J$	$Q_1\Box\!\!\rightarrow R$	$\cdots$	$Q_J\Box\!\!\rightarrow R$	$P\Box\!\!\rightarrow R$	$\sim P\Box\!\!\rightarrow R$
1	T	T	$\cdots$	T	T	T	$\cdots$	T	T	T
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$