Empir Econ
DOI 10.1007/s00181-016-1186-1

Assessing the evidence on neighborhood effects from Moving to Opportunity

Dionissi Aliprantis[1]

Received: 15 December 2015 / Accepted: 10 August 2016

Abstract The Moving to Opportunity (MTO) experiment randomly assigned housing vouchers that could be used in low-poverty neighborhoods. Consistent with the literature, I find that receiving an MTO voucher had no effect on outcomes like earnings, employment, and test scores. However, after studying the assumptions identifying neighborhood effects with MTO data, this paper reaches a very different interpretation of these results than found in the literature. I first specify a model in which the absence of effects from the MTO program implies an absence of neighborhood effects. I present theory and evidence against two key assumptions of this model: that poverty is the only determinant of neighborhood quality and that outcomes only change across one threshold of neighborhood quality. I then show that in a more realistic model of neighborhood effects that relaxes these assumptions, the absence of effects from the MTO program is perfectly compatible with the presence of neighborhood effects. This analysis illustrates why the implicit identification strategies used in the literature on MTO can be misleading.

Keywords Moving to Opportunity · Neighborhood effect · Program effect

JEL Classification C30 · H50 · I38 · J10 · R00

1 Introduction

Understanding neighborhood effects is an imperative for public policy. For example, debates about the role of government in education cannot be resolved without

✉ Dionissi Aliprantis
dionissi.aliprantis@clev.frb.org

1 Research Department, Federal Reserve Bank of Cleveland, PO Box 6387, Cleveland, OH 44101-1387, USA

Published online: 16 January 2017

understanding the nature of effects from localized differences in resources and social interactions (Friedman 1955; Manski 2013b). Likewise, empirically characterizing neighborhood effects is crucial for understanding the persistence of racial inequality in the United States, and for designing effective policy in response (Wilson 1987; Sampson 2012).

Conclusive evidence on neighborhood effects is elusive, though, since spatial correlations in outcomes could reflect residential sorting as easily as neighborhood effects. To overcome this fundamental selection issue, researchers have studied housing mobility programs like Gautreaux, which relocated 7100 public housing families throughout Chicago in a quasi-random manner between 1976 and 1998 (Polikoff 2006). The results from Gautreaux have been interpreted as strong evidence of neighborhood effects: 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 (Rubinowitz and Rosenbaum 2000; Rosenbaum 1995; Mendenhall et al. 2006).

The Moving to Opportunity (MTO) housing mobility program was designed to replicate the success of Gautreaux by randomly allocating housing vouchers to public housing residents in five US cities between 1994 and 1998. In a tremendous disappointment, the results from the MTO program were not as positive as the results from the Gautreaux program. There were no statistically significant improvements in education and labor market outcomes (Sanbonmatsu et al. 2006; Kling et al. 2007a), and the risky behavior of young males actually grew worse (Kling et al. 2005).

The majority of the literature has interpreted the results from MTO as evidence against neighborhood effects. For example, Ludwig et al. (2013) interpret the results from the MTO program as being "Contrary to the widespread view that living in a disadvantaged inner-city neighborhood depresses labor market outcomes, ..." (p. 228). Angrist and Pischke (2010)'s interpretation of MTO is that "The program has produced surprising and influential evidence weighing against the view that neighborhood effects are a primary determinant of low earnings by the residents of poor neighborhoods" (p. 4).

Interpreting MTO as evidence against neighborhood effects has previously come under criticism for conflating program effects with neighborhood effects (Clampet-Lundquist and Massey 2008). However, this critique has been dismissed as reflecting a misunderstanding of selection bias (Ludwig et al. 2008). The literature continues to interpret MTO as an experiment that randomly allocated households to varying peer environments because housing vouchers were randomly assigned (Angrist 2014).

This paper shows that the distinction made in Clampet-Lundquist and Massey (2008) between program effects and neighborhood effects is in fact critical to assessing the evidence on neighborhood effects from MTO. To make the issues clear, one must first consider a standard joint model of potential outcomes and selection into treatment and note the following: Defining treatment as moving with an MTO voucher generates a model of program effects, while defining treatment as moving to a high-quality neighborhood generates a model of neighborhood effects.

I ask a question that follows from Clampet-Lundquist and Massey (2008)'s analysis: What model of neighborhood effects can be used to justify the view in the literature that "If neighborhood environments affect behavior ... then these neighborhood effects ought to be reflected in ITT [Intent-to-Treat] and TOT [Treatment-on-the-Treated] impacts [of the program] on behavior" (Ludwig et al. 2008, pp. 181–182)? By investigating this question, I not only distinguish between program and neighborhood effects, but also establish assumptions about models of neighborhood effects under which researchers can use program effects to learn about neighborhood effects. I find that these assumptions are strong, have led the literature to draw unwarranted conclusions from the MTO results, and can be relaxed by directly estimating a neighborhood effects model.

Put a bit more precisely, suppose that Y is an outcome variable like employment, D is neighborhood quality, Z is receipt of a housing voucher, and consider a model of neighborhood effects consisting of potential outcomes $Y(D)$ and $D(Z)$. Randomization of a housing voucher $Z \in \{0, 1\}$ identifies a class of program effects, the potential outcomes $Y(Z)$ and $D(Z)$. A central contribution of Clampet-Lundquist and Massey (2008) was to make a distinction between program effects $Y(Z)$ and neighborhood effects $Y(D)$.

This paper asks the further question: What definition of D and resulting assumptions about $Y(D)$ allow us to draw conclusions about neighborhood effects from program effects? Different specifications of D, such as $D \in \{0, 1\}$ or $D \in \{1, 2, \ldots, J\}$, generate different models. These different models make distinct assumptions about how changing neighborhood characteristics affect outcomes. I show two sufficient assumptions for learning about neighborhood effects from program effects are that neighborhood quality is a binary variable and that poverty is a proxy for quality. The resulting specification of potential outcomes $Y(D)$ imposes that the outcome variable Y changes only in response to crossing a single threshold of neighborhood poverty.

In more general models of neighborhood effects that relax these assumptions, it is entirely possible that neighborhood environments affect behavior but that these neighborhood effects are not reflected in the effects of the MTO program. I provide empirical evidence and theoretical arguments in favor of adopting a more general model of this type. I first show that outcomes in the model should be allowed to change across more than just one margin of quality. I then show that in order to test Wilson's theory of neighborhood effects, neighborhood quality should be defined as a function of other characteristics in addition to poverty. In order to conduct my empirical analysis, I use principal components analysis to construct a scalar measure of neighborhood quality that is a function of not only the neighborhood poverty rate, but also the percent with high school degrees, the percent with BAs, the percent of single-headed households, the male employment-to-population ratio, and the female unemployment rate.[1]

I first provide theory and evidence in favor of adopting a model with more than two levels of quality. I show that MTO only induced transitions across low levels of neighborhood quality. As a result, MTO did not generate the variation in neighborhood quality necessary to learn whether changes to many types of neighborhood environments would alter outcomes. In other words, the neighborhood effects model with

[1] My measure of quality is a normalization of the first principal component of these variables, or the one-dimensional vector explaining the most variation in these variables.

only two levels of quality implicitly used in the literature on MTO simply assumes that changes to many types of neighborhood environments would not alter outcomes.

I also provide empirical evidence against using poverty as a proxy for quality in the MTO experiment. I show that there are many low-poverty neighborhoods in MTO states that are still low quality. Thus, even when focused on understanding effects from moves across low levels of quality, researchers must still be careful to identify the moves induced by MTO that changed neighborhood quality in conjunction with neighborhood poverty.

The paper proceeds as follows: Sect. 2 describes the MTO experiment. Section 3 characterizes the current literature on MTO in terms of the neighborhood effects model assumptions it implicitly imposes. Subsequent sections present theoretical reasoning and empirical evidence on these assumptions and how they might be relaxed. Section 4 presents a canonical joint model of potential outcomes and selection into treatment without any view of how such a model might be applied to MTO. Sections 5.1 and 5.2 then proceed, respectively, to discuss the program and neighborhood effects identified with the MTO data set under various assumptions. Section 6 concludes.

2 Moving to Opportunity (MTO)

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 <30% 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).[2]

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

[2] 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).

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 $<10\%$. However, after 1 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 nonprofit. 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.[3]

3 What model of neighborhood effects can justify the literature's current interpretation of MTO?

Program effects and neighborhood effects are different parameters defined in distinct models (Heckman 2010). Yet intent-to-treat (ITT) and treatment-on-the-treated (TOT) effects from receiving an MTO voucher have been interpreted as evidence on neighborhood effects in the literature on MTO. For example, Kling et al. (2007a) include ITT and TOT program effect estimates as "direct evidence on the existence, direction, and magnitude of neighborhood effects" (p. 84), and Ludwig et al. (2008) contend that "Both [ITT and TOT] estimators are informative about the existence of neighborhood effects on behavior" (p. 146).

What model of neighborhood effects can justify these statements? The current interpretation of the results from MTO does not equate program and neighborhood effects, but rather combines evidence on program effects from MTO together with logical arguments to indirectly draw conclusions about neighborhood effects.[4] This section shows that such an interpretation of MTO relies on an implicit, and therefore poorly specified, model of neighborhood effects.

Suppose we were only focused on comparing the MTO experimental and control groups, and that for the sake of exposition we are focused on the single outcome of adult employment. The focus on adult employment is motivated by the fact that conclusions about neighborhood effects on this outcome have been reached based on the lack of large treatment effects from the MTO program.[5] The following statement:

[3] 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.

[4] This is the author's current interpretation of the literature, most prominently represented by Kling et al. (2007a) and Ludwig et al. (2008). However, the distinction between program and neighborhood effect parameters has not always been made clearly. 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 Ludwig et al. (2013), Sanbonmatsu et al. (2012), and Gennetian et al. (2012).

[5] This interpretation of the results from MTO can be found in Kling et al. (2007a), Ludwig et al. (2013, pp. 228–229), Angrist (2014, p. 106), Angrist and Pischke (2010, p. 4). Some preliminary instrumental variable analysis can be found in Ludwig et al. (2008), and recent papers like Aliprantis and Richter (2016) and

(†): "If neighborhood environments affect behavior... then these neighborhood effects ought to be reflected in ITT and TOT impacts [of the program] on behavior" (Ludwig et al. 2008, pp. 181–182).

can be justified by a model of potential outcomes $D(Z)$, $Y(D)$, and $Y(Z)$ under the assumptions that D is a binary indicator of neighborhood quality, Z is a binary indicator of receiving an MTO voucher versus being in the control group, and Y is a binary indicator of employment:

M1: $D_i \equiv \mathbf{1}\{\text{individual } i \text{ lives in a high-quality neighborhood}\}$
M2: $Z_i \equiv \mathbf{1}\{\text{individual } i \text{ received an MTO voucher}\}$
M3: $Y_i \equiv \mathbf{1}\{\text{individual } i \text{ is employed}\}$

Note that treatment is defined here in terms of neighborhood quality, whereas most of the literature on MTO estimates models in which treatment is defined as moving with an MTO voucher.[6] It is important to distinguish between these definitions of treatment because they generate distinct models of potential outcomes and selection, with one being a model of program effects (D1), and the other being a model of neighborhood effects (D2):

D1 Treatment is moving with the aid of the program (i.e., using an MTO voucher).
D2 Treatment is moving to a high-quality neighborhood.

Without any further empirical or theoretical restrictions on maintained assumptions M1–M3, these variables result in a neighborhood effects model that can generate any of $4^3 = 64$ possible counterfactual worlds displayed in Table 7, found in Appendix 1. In terms of the analysis of treatment response (Manski 2011), one could think of these counterfactual worlds as representing the average individual response functions for various states of the world.

To gain some intuition about the possible states of the world (shown in full as Table 7 in "Appendix 1"), consider States 22 and 32 as shown in Table 1. In both States 22 and 32 there are program effects on individual i's neighborhood quality since $D(Z = 1) = 1$ and $D(Z = 0) = 0$. Columns 1 and 2 indicate that the individual would move to a "good" neighborhood when receiving a voucher, but would remain in a low-quality neighborhood without a voucher.

States 22 and 32 differ according to the presence of program and neighborhood effects on individual i's employment. Columns 5 and 6 indicate that in State 22 the individual would have a job with a voucher $[Y(Z = 1) = 1]$, but not without a voucher $[Y(Z = 0) = 0]$. And columns 3 and 4 indicate that in State 22 individual i would have a job when living in a "good" neighborhood $[Y(D = 1) = 1]$, but not when in a "bad" neighborhood $[Y(D = 0) = 0]$. In contrast, State 32 is characterized by no program effects on employment [columns 5 and 6 showing $Y(Z = 1) = Y(Z = 0) = 0$] and no neighborhood effects on employment [columns 3 and 4 showing $Y(D = 1) = Y(D = 0) = 0$].

Footenote 5 Continued
Pinto (2014) that have estimated neighborhood effects models using the MTO data have found evidence of neighborhood effects on adult employment.

[6] See the Appendix of Ludwig et al. (2008) or Ludwig et al. (2013) for examples.

Table 1 Some states of the world possible in an unrestricted neighborhood effects model with binary variables

Column	1	2	3	4	5	6
Row	$D(Z=1)$	$D(Z=0)$	$Y(D=1)$	$Y(D=0)$	$Y(Z=1)$	$Y(Z=0)$
(State 22)	1	0	1	0	1	0
(State 32)	1	0	0	0	0	0

$Z \equiv$ individual i receives an MTO voucher, $D \equiv$ individual i lives in a ‘good’ neighborhood, $Y \equiv$ individual i is employed

One could combine theory and empirical observations to rule out that the state of the world as observed in the MTO data looked like some of the possible states of the world in Table 7 of “Appendix 1”. For example, based on empirical observations from MTO on the neighborhoods of residence of control group households as recorded at the time of the follow-up survey, it is likely to be uncontroversial that we can rule out $D(Z = 0) = 1$, or living in a “good” neighborhood without a voucher, in the real world. This would eliminate States 1–16 or 33–48 from representing the real world, leaving the 32 states of worlds displayed in Table 2 in consideration for accurately describing the world as observed in MTO.

So far this approach to relating program effects and neighborhood effects has only used empirical observations in addition to binary definitions of variables to rule out states of the world. It would be possible to further rule out from consideration some of the states from Table 2 solely on the basis of theory. One possibility would be to adopt the neighborhood effects model shown in Fig. 1, along with the new model of neighborhood effects resulting from the MTO intervention.

One could apply this neighborhood effects model to rule out particular states of world from consideration. For example, this would rule out States 18, 19, and 20 as simply being inconsistent with the types of counterfactuals believed to be similar to those in the current state of the world, as expressed by the restrictions on the Data Generating Process placed by the model.[7] One could proceed to eliminate states of the world from Table 2, with the states dropped all following the same pattern of elimination: They either contradict empirical observation, require that the MTO voucher affects outcomes through some pathway other than neighborhood quality, or else would require some column to take different values in order to be consistent with our model.

Suppose that Table 3 does in fact represent the states of the world that could possibly correspond with the true state of the world *under the assumptions* of the model (Fig. 1 and D2) and M1–M3. Under these assumptions, and a few more, one can use evidence on the program effects pertaining to $D(Z)$ and $Y(Z)$ to draw conclusions about the

[7] State 18 describes a state of the world in which an individual will be employed regardless of the neighborhood in which they reside, yet receiving an MTO voucher will cause them to become employed. State 19 implies that an individual will be employed regardless of the neighborhood in which they reside, yet receiving an MTO voucher will cause them to become *unemployed*. Finally, State 20 describes a state of the world in which the individual is both always employed (columns 3 and 4) or else is never employed (columns 5 and 6), which simply cannot happen in our model as structured.

Table 2 States of the world possible in empirically-restricted neighborhood effects model with binary variables

Column	1	2	3	4	5	6
Row	$D(Z=1)$	$D(Z=0)$	$Y(D=1)$	$Y(D=0)$	$Y(Z=1)$	$Y(Z=0)$
After restrictions imposed by empirical observations						
(State 17)	1	0	1	1	1	1
(State 18)					1	0
(State 19)					0	1
(State 20)					0	0
(State 21)			1	0	1	1
(State 22)					1	0
(State 23)					0	1
(State 24)					0	0
(State 25)			0	1	1	1
(State 26)					1	0
(State 27)					0	1
(State 28)					0	0
(State 29)			0	0	1	1
(State 30)					1	0
(State 31)					0	1
(State 32)					0	0
(State 49)	0	0	1	1	1	1
(State 50)					1	0
(State 51)					0	1
(State 52)					0	0
(State 53)			1	0	1	1
(State 54)					1	0
(State 55)					0	1
(State 56)					0	0
(State 57)			0	1	1	1
(State 58)					1	0
(State 59)					0	1
(State 60)					0	0
(State 61)			0	0	1	1
(State 62)					1	0
(State 63)					0	1
(State 64)					0	0

$Z \equiv$ individual i receives an MTO voucher, $D \equiv$ individual i lives in a "good" neighborhood, $Y \equiv$ individual i is employed

neighborhood effects represented by $Y(D)$. To begin, since Z is randomized one can learn about $D(Z)$ and $Y(Z)$ from the values of $E[D|Z]$ and $E[Y|Z]$ observed in MTO.

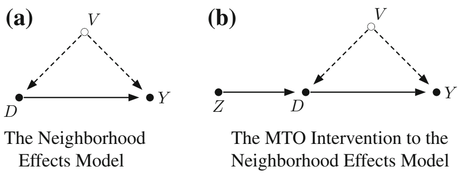

Fig. 1 Directed acyclic graphs of the neighborhood effects model. Note: This figure follows the convention from Pearl (2009) of communicating that a variable is observed by drawing a solid line to its descendants, and communicating that a variable is unobserved by drawing a dashed line to its descendants. These models correspond to the neighborhood effects model in Sect. 4 under assumptions A1–A6, definition of treatment D2, and V in the figure defined to be (U_D, U_0, U_1)

Table 3 States of the world possible in empirically and theoretically restricted neighborhood effects model

Column	1	2	3	4	5	6
Row	$D(Z=1)$	$D(Z=0)$	$Y(D=1)$	$Y(D=0)$	$Y(Z=1)$	$Y(Z=0)$
After restrictions imposed by empirical observations and theory (i.e., the model)						
(State 17)	1	0	1	1	1	1
(State 22)			1	0	1	0
(State 27)			0	1	0	1
(State 32)			0	0	0	0
(State 49)	0	0	1	1	1	1
(State 56)			1	0	0	0
(State 57)			0	1	1	1
(State 64)			0	0	0	0

$Z \equiv$ individual i receives an MTO voucher, $D \equiv$ individual i lives in a "good" neighborhood, $Y \equiv$ individual i is employed

If one also adopts the assumptions:

NQB Neighborhood quality D is a binary function of a latent index of neighborhood quality q: $D \equiv \mathbf{1}\{q \geq q^*\}$
NQP Neighborhood quality q is a one-dimensional vector that is a scalar function of neighborhood poverty p: $q = \alpha p$

then the reasoning proceeds that the changes in neighborhood poverty observed in MTO imply that the true state of the world must be in one of States 17, 22, 27, or 32. Within these states, only 22 and 27 "exhibit neighborhood effects" (see columns 3 and 4), and in these states there are also program effects (see columns 5 and 6). Thus, *under the adopted modeling assumptions*, the empirical evidence can justify statement (†).

Once statement (†) is justified, conclusions about neighborhood effects follow quickly. The reasoning proceeds looking at columns 5 and 6. The empirical evidence on program effects indicates that the true state of the world is either States 32, 56, or 64. Combined with the observed changes in neighborhood poverty rates implying the true state is in one of States 17, 22, 27, or 32, the true state of the world must be State 32. Thus, one concludes:

(⋆): The evidence from MTO suggests neighborhood effects are not strong.

Because statement (†) is false in more general models of neighborhood effects relaxing assumptions NQB and NQP, conclusion (⋆) need not be true in such models.[8] I now consider theoretical and empirical evidence in favor of relaxing assumptions NQB and NQP.

4 The definition of causal effects

4.1 A joint model of potential outcomes and selection

I now define several treatment effect parameters within a standard model of potential outcomes and selection into treatment (Heckman and Vytlacil 2005; Imbens and Rubin 2015), initially taking no stand on what effects the researcher aims to identify. 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

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

The measured outcome variable Y is

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

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

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

Note that these are not structural equations under Definition 5.4.1 in Pearl (2009).[9] Thus, since unobserved factors U_0 and U_1 influence $Y(0)$ and $Y(1)$, respectively, exclusion restrictions will need to be made if particular variables are to be ruled out of being a part of U_0 or U_1.

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

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

[8] Aliprantis and Richter (2016) is one example of neighborhood effects estimated under weaker assumptions than NQB and NQP in which the estimated effects contradict conclusion (⋆).

[9] See Aliprantis (2015a, b) or Heckman and Vytlacil (2005) for further discussion.

Noting it need not be true that $D = Z$, $D(Z)$ will denote the treatment received when assigned treatment Z and there is an explicit model of how individuals select into treatment. 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 $\pi^Z(X) \equiv F_V(\mu_D(X, Z)) \equiv Pr(D = 1|X, Z)$.

I adopt a simple version of Heckman and Vytlacil (2005) and Heckman et al. (2006) by assuming:

A1 $\gamma_i = \gamma$ for all i and $\gamma \neq 0$
A2 $\{U_j, V\} \perp\!\!\!\perp Z \mid X$ for $j = 0, 1$
A3 The distribution of V is absolutely continuous
A4 $E\big[|Y(0)|\big|X\big] < \infty$ and $E\big[|Y(1)|\big|X\big] < \infty$
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 one might be interested in investigating. It is standard to define the ITT, TOT, and local average treatment effect (LATE) parameters as follows:

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

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

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

Note that so far no assumption has been made about the relationship between the unobservable components determining potential outcomes and selection into treatment. The treatment effects defined in Eqs. 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 relationship between the unobservables in the model. Two mutually exclusive (but not exhaustive) assumptions often adopted in the literature are Ignorability and Essential Heterogeneity:

Ig $\{U_1, U_0\} \perp\!\!\!\perp V|\ X$.
EH $\text{COV}(U_1 - U_0, V) \mid X \neq 0$.

5 The identification of causal effects

5.1 What program effects are identified by MTO?

Since the model defined in Sect. 4.1 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. I 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 one might wish to consider is:

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

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

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

Consider an additional restriction placed on the choice model,

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

Under A5*

$$D(1) - D(0) = 1 \Longleftrightarrow D(1) = 1, \tag{10}$$

and thus under either assumptions (A1–A6, Ig, D1) or assumptions (A1–A6, A5*, Ig, D1) the Wald estimator allows one to identify the homogeneous program effect of MTO:

$$\frac{E[Y|x, Z=1] - E[Y|x, Z=0]}{E[D|x, Z=1] - E[D|x, Z=0]} = \triangle^{\mathrm{TOT}}(x) = \triangle^{\mathrm{LATE}}(x, \cdot, \cdot) \tag{11}$$

If one relaxes Ig by assuming EH, then under (A1–A6, EH, D1) MTO identifies the following program effect that is determined in part by selection into treatment:

$$\frac{E[Y|x, Z=1] - E[Y|x, Z=0]}{E[D|x, Z=1] - E[D|x, Z=0]} = \triangle^{\mathrm{LATE}}\left(x, \pi^0(x), \pi^1(x)\right). \tag{12}$$

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

$$\frac{E[Y|x, Z=1] - E[Y|x, Z=0]}{E[D|x, Z=1] - E[D|x, Z=0]} = \triangle^{\mathrm{TOT}}(x) = \triangle^{\mathrm{LATE}}\left(x, 0, \pi^1(x)\right). \tag{13}$$

Since assumptions (A1–A6, A5*, EH, D1) appear reasonable together, the program effect in Eq. 13 is identified by MTO. “Appendix 3” has a further discussion of assumptions about the distribution of unobserved variables, and “Appendix 4” a discussion of the external validity of this parameter.

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 5–7 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 some outcomes of young males (Kling et al. 2005, 2007a).

5.2 What neighborhood effects are identified by MTO?

Another treatment whose effects one 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 Eqs. 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.[10] 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.

5.2.1 Defining neighborhood quality and assumption A2

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: $D \equiv \mathbf{1}\{q \geq q^*\}$

[10] While using an MTO voucher did initially require moving to a neighborhood with particular poverty characteristics ($<10\%$), this requirement only had to be met for 1 year. Since subsequent moves were frequent, often involuntary, and tended to be to low-quality neighborhoods (de Souza Briggs et al. 2010; Sampson 2008), the initial MTO move does not to capture the entire sequence of neighborhood characteristics, even when measured by poverty alone. Here I measure mobility using residence at the time of the interim evaluation, but other ways of dealing with dynamics, whether within the static models discussed here or within an expanded dynamic model, could also be appropriate.

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 (i.e., $q^* = 80$). 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_1 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.[11]

One way to resolve this issue is to generalize the model in Sect. 4.1 in terms of the ordered choice model developed in Heckman et al. (2006).[12] A generalized framework assumes

NQJ Neighborhood quality D is a multi-valued function of a latent index of neighborhood quality q: $D \equiv j \times \mathbf{1}\{C_{j-1} < q \leq C_j\}$ 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.

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 effects are constant across unobservables.[13]

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-poverty ($D = j - 1$) 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 $Z = 1$.

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

[12] An alternative and complementary approach is to use an unordered choice model as in Pinto (2014).

[13] 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. Effects in these analyses are constant in U under the specification in Eq. 3 since they assume $U_j = U$ for all $j \in \{1, \ldots, J\}$, so $U_{j+1,i} - U_{j,i} = U_i - U_i = 0$.

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 Sect. 5.2.2.

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.

5.2.2 Empirical evidence on assumptions A5, NQP, and NQK

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 employment-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.[14]

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

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

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 Sect. 5.2.1, I 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

[14] 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 time trends. Second, sampling weights must be used to account for the subsampling procedures used during the interim evaluation data collection.

Table 4 Principal components analysis

Coefficients on first eigenvector		Proportion of variance explained		
Variable	Coefficient	Eigenvector	Eigenvalue	Proportion
Poverty rate	−0.45	1	4.14	0.69
HS graduation rate	0.44	2	0.67	0.11
BA attainment rate	0.40	3	0.51	0.08
Percent single-headed HHs	−0.36	4	0.35	0.06
Male EPR	0.41	5	0.22	0.04
Female unemployment rate	−0.39	6	0.12	0.02

This table reports the results of principal components analysis conducted on decennial US Census data from 2000 using the national percentiles (in terms of population) of census tract poverty rate, high school graduation rate, BA attainment rate, share of single-headed households, the male employment-to-population ratio, and the female unemployment rate

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.

There are several variables not included in the index. I do not include race in the index for the same reason I do not include eye color: My theory of race is that it is a set of physical characteristics distributed independently of the distribution of potential outcomes in a model of neighborhood effects.[15] I also do not include a measure of house prices like median rent. This is to ensure that the variables in the index all have clear interpretations in terms of mechanisms described in Wilson (1987), and because an index excluding median rent explains 99% of the variation of an index including median rent. Finally, I do not include physical characteristics that could be important, such as public transportation, green spaces, or access to supermarkets, because these variables are harder to obtain and appropriately measure.

Table 4 shows that the univariate index resulting from principal components analysis explains 69% of the variance of these neighborhood characteristics and that no additional eigenvector would explain more than 11% of the variance of these variables. Table 4 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 2a shows the expected negative correlation between neighborhood quality and neighborhood poverty rate. One can see in Fig. 2b that the US population distribution of neighborhood poverty rates in 2000 had a long right tail. Similarly, Fig. 2c shows that the US population distribution of neighborhood quality had a long left tail in 2000. Figure 2d, e shows how far in the tails of these national distributions much of the MTO sample typically resided.

[15] Nevertheless, race will be correlated with the neighborhood characteristics causally affecting outcomes due to the history of racial discrimination in the USA. Aliprantis and Kolliner (2015) study race and neighborhood characteristics in the context of MTO.

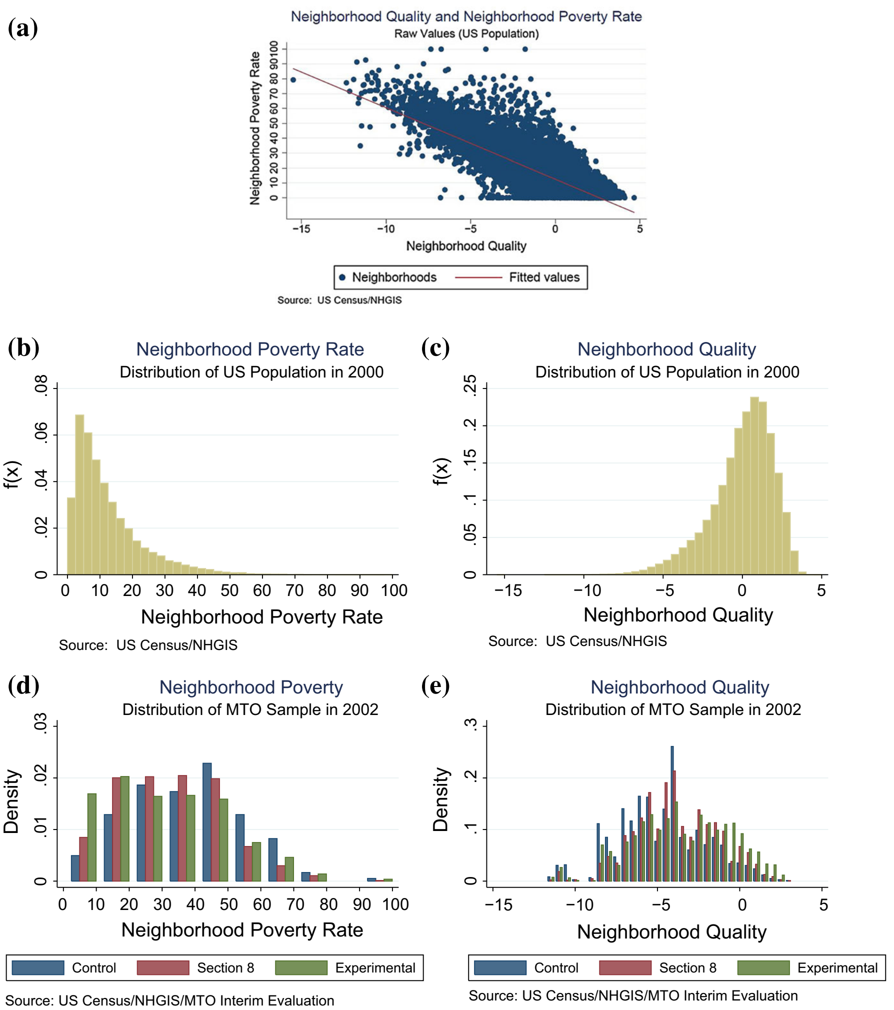

Fig. 2 Neighborhood poverty rate and neighborhood quality. This figure shows the distribution of quality as obtained from principal components analysis conducted on decennial US Census data from 2000 as detailed in the text, as well as the national percentile (in terms of population) of the 2000 US census tract poverty rate. **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

Moving from a neighborhood with a poverty rate of 70% to a neighborhood with a 50% poverty rate might be a large change in the poverty rate, but Fig. 2b suggests that one should also consider how big this change is relative to the national distribution of neighborhoods. 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

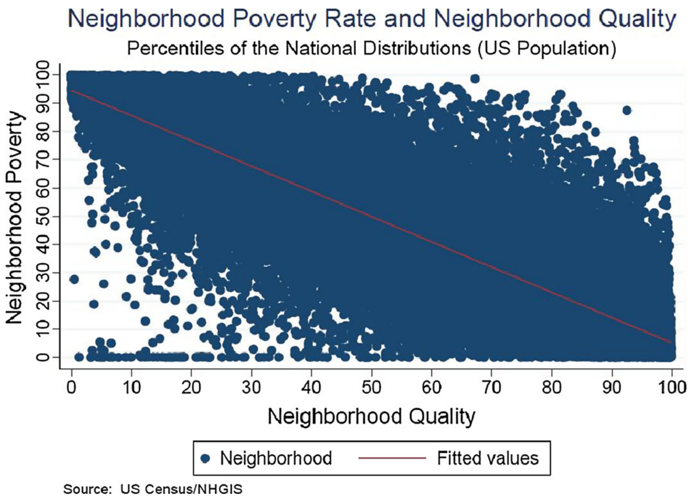

Fig. 3 Neighborhood poverty and quality. This figure shows a scatterplot of percentiles of census tract poverty rate on the y-axis and percentiles of census tract quality on the x-axis. Both percentiles pertain to the national distribution of the US population in 2000

Table 5 Low-poverty ($\leq$10%), low-quality ($D \leq 3$) neighborhoods in MTO states in 2000

Nbd quality	Number of residents
$D = 1$	6362
$D = 2$	93,385
$D = 3$	751,738

This table reports the existence of low-quality census tracts that met the experimental MTO cutoff by having a 10% poverty rate or less

Fig. 3. This figure shows that although the expected negative relationship still remains, there is a considerable range for 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 range may not be surprising given the coefficients reported in Table 4, and can also be seen in Table 5, which presents evidence that in MTO states there were many low-poverty neighborhoods that were also in the second and third deciles of the national distribution of quality. While the empirical evidence supports the adoption of assumption NQK over NQP if neighborhood characteristics other than poverty influence outcomes, simply comparing assumptions NQK and NQP in a theoretical way highlights that even defining neighborhood quality requires explicitly specifying which neighborhood characteristics influence outcomes.

Figure 4 shows that very few MTO adults were induced into high-quality neighborhoods.[16] At the time of the interim evaluation <10% of the experimental group lived

[16] It is worth noting that the same general conclusion also holds in models assuming NQP. For example, Quigley and Raphael (2008) point out that “The effect of treatment under the MTO program was, on average,

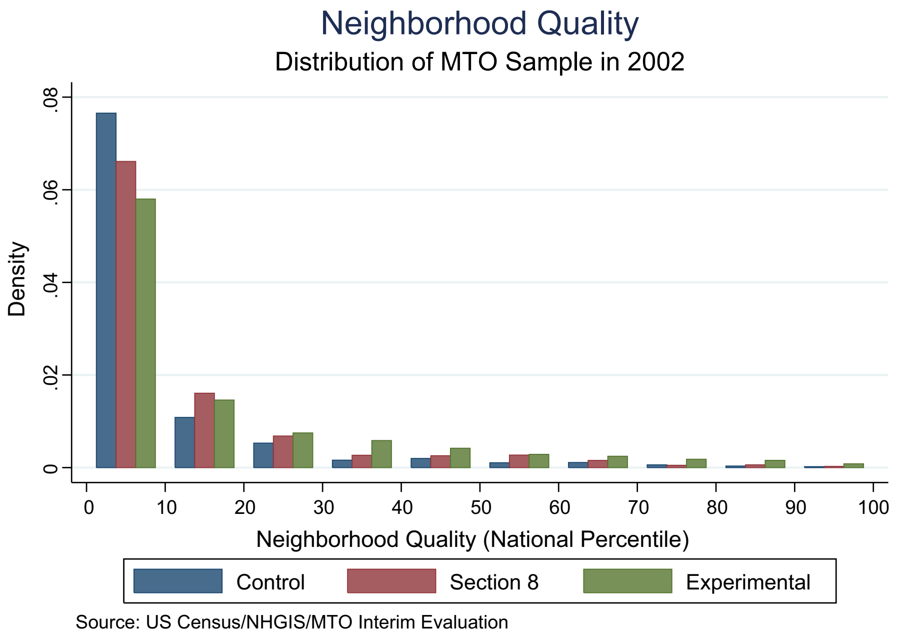

Fig. 4 Neighborhood quality of MTO participants in 2002. This figure shows the distribution of MTO participants at the time of the interim evaluation survey according to the index of quality discussed in the text, measured in percentiles of the national distribution of the US population in 2000

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.[17]

The distributions in Fig. 4 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 $<20\%$ of the MTO experimental group lived in neighborhoods above the 30th percentile of the national distribution of quality, and $<10\%$ lived in neighborhoods above the median.

To get a sense of the changes induced by MTO in specific neighborhood characteristics, consider the share of compliers when neighborhood quality is a one-dimensional

Footnote 16 continued
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).

[17] DeLuca and Rosenbaum (2003) find that 66% of the suburban group and 13% of the city group lived in the suburbs of Chicago 14 years after original placement through Gautreaux. DeLuca and Rosenbaum (2003) cite limited availability of housing, rather than the choice to not move through the program, as the reason only 20% of eligible applicants moved through Gautreaux. This claim is based on evidence that 95% 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.

Table 6 Share of compliers for various binary definitions of quality

$E[D_i \mid Z_i = 1] - E[D_i \mid Z_i = 0]$ where $D_i = \mathbf{1}\{q_i \geq \text{percentile}\}$		
Neighborhood variable	25th percentile	50th percentile
BA attainment rate	0.16	0.09
Poverty rate	0.16	0.08
HS graduation rate	0.16	0.07
Quality	0.13	0.07
Female unemployment rate	0.11	0.05
Male EPR	0.10	0.06
Percent single-headed HHs	0.07	0.04

binary variable defined as in **NQB** in terms of the 25th and 50th percentile of the US population distribution in 2000, and when the instrument is receiving either an experimental voucher or a control group subsidy. Consistent with the evidence in Fig. 4, Table 6 shows that MTO induced $<10\%$ of households into above-median neighborhoods along any of the characteristics considered. The largest changes in neighborhood characteristics induced by MTO were in terms of educational attainment and poverty, and the smallest changes were in terms of labor market outcomes and the share of single-headed households.

6 Conclusion

Should Moving to Opportunity be interpreted as a test of Wilson (1987)'s model of neighborhood effects? One prominent group of researchers interprets the results from MTO in this way:

> In Wilson's model, the exodus of middle- and working-class families was particularly important because these families served as "a social buffer," as "mainstream role models that help keep alive the perception that education is meaningful, that steady employment is a viable alternative to welfare, and that family stability is the norm, not the exception" (Wilson 1987, p. 49). MTO as implemented would seem to provide an almost perfect test of this theory–it helped families move out of some of the most unsafe neighborhoods in America into neighborhoods with substantial shares of middle-class minority residents who could potentially serve as role models (Ludwig et al. 2008, p. 163).

This paper presented evidence that such a view over-interprets the results from MTO. MTO did not move a large share of families into neighborhoods with substantial shares of residents with high school diplomas, college degrees, where the male employment-to-population ratio was high, the female employment rate was high, and in which there were few single-headed households. As a result, interpreting the effects of MTO as a test of the existence of neighborhood effects requires adopting a model of neighborhood effects with strong assumptions that would be avoided if stated explicitly.

Acknowledgements I thank Francisca G.-C. Richter, Jeffrey Kling, my Math Corps students, and several seminar participants and anonymous referees for contributing to this paper. I am also grateful to Mary Zenker for research assistance and Paul Joice at HUD for help accessing the data. The research reported here was supported in part by the Institute of Education Sciences, US Department of Education, through Grant R305C050041-05 to the University of Pennsylvania. The views stated herein are those of the author and are not necessarily those of the Federal Reserve Bank of Cleveland, the Board of Governors of the Federal Reserve System, or the US Department of Education.

Appendix 1: Full contingency table for states of world

See Table 7.

Table 7 States of the world possible in unrestricted neighborhood effects model with binary variables

Column	1	2	3	4	5	6
Row	$D(Z=1)$	$D(Z=0)$	$Y(D=1)$	$Y(D=0)$	$Y(Z=1)$	$Y(Z=0)$
(State 1)	1	1	1	1	1	1
(State 2)					1	0
(State 3)					0	1
(State 4)					0	0
(State 5)			1	0	1	1
(State 6)					1	0
(State 7)					0	1
(State 8)					0	0
(State 9)			0	1	1	1
(State 10)					1	0
(State 11)					0	1
(State 12)					0	0
(State 13)			0	0	1	1
(State 14)					1	0
(State 15)					0	1
(State 16)					0	0
(State 17)	1	0	1	1	1	1
(State 18)					1	0
(State 19)					0	1
(State 20)					0	0
(State 21)			1	0	1	1
(State 22)					1	0
(State 23)					0	1
(State 24)					0	0
(State 25)			0	1	1	1
(State 26)					1	0
(State 27)					0	1
(State 28)					0	0

Table 7 continued

Column	1	2	3	4	5	6
Row	$D(Z=1)$	$D(Z=0)$	$Y(D=1)$	$Y(D=0)$	$Y(Z=1)$	$Y(Z=0)$
(State 29)			0	0	1	1
(State 30)					1	0
(State 31)					0	1
(State 32)					0	0
(State 33)	0	1	1	1	1	1
(State 34)					1	0
(State 35)					0	1
(State 36)					0	0
(State 37)			1	0	1	1
(State 38)					1	0
(State 39)					0	1
(State 40)					0	0
(State 41)			0	1	1	1
(State 42)					1	0
(State 43)					0	1
(State 44)					0	0
(State 45)			0	0	1	1
(State 46)					1	0
(State 47)					0	1
(State 48)					0	0
(State 49)	0	0	1	1	1	1
(State 50)					1	0
(State 51)					0	1
(State 52)					0	0
(State 53)			1	0	1	1
(State 54)					1	0
(State 55)					0	1
(State 56)					0	0
(State 57)			0	1	1	1
(State 58)					1	0
(State 59)					0	1
(State 60)					0	0
(State 61)			0	0	1	1
(State 62)					1	0
(State 63)					0	1
(State 64)					0	0

$Z \equiv$ individual i receives an MTO voucher, $D \equiv$ individual i lives in a "good" neighborhood, $Y \equiv$ individual i is employed

Appendix 2: The neighborhood effects identified by MTO

Effects from moving to high-quality neighborhoods are not identified by MTO. Given the evidence in Sect. 5.2.2, 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, EH, D2-NQB) the Wald estimator identifies the LATE:

$$\frac{E[Y|x, Z=1] - E[Y|x, Z=0]}{E[D|x, Z=1] - E[D|x, Z=0]} = \triangle^{\text{LATE}} \left(x, \pi^0(x), \pi^1(x)\right) \tag{14}$$

If we believe assumption A2 will fail to hold when treatment is defined under D2-NQB for the reasons discussed in Sect. 5.2.1, we could alternatively define treatment under D2-NQJ to generate a transition-specific analogue to 14:

$$\begin{aligned}&\triangle^{\text{LATE}}_{j,j+1} \left(x, \pi^0_j(x), \pi^1_j(x)\right)\\ &\quad \equiv E\left[Y(D=j+1) - Y(D=j)|x, D(Z=1)=j+1, D(Z=0)=j\right].\end{aligned}$$

Versions of the model have been estimated in Kling et al. (2007a) and Ludwig et al. (2008) under (A1–A6, SI, and D2-NQJ-NQP). A dose–response analysis is used in Kling et al. (2007a) to determine whether parameters are constant across all j to $j+1$ transitions in $\{1, \ldots, J\}$. Aliprantis and Richter (2016) estimate the model under (A1–A6, EH, D2-NQJ-NQK). That analysis makes A2 more plausible by relaxing D2-NQJ-NQP–D2-NQJ-NQK and allows for the identification and estimation of LATEs that are heterogeneous over unobservables by relaxing SI to EH.[18]

Appendix 3: Assumptions about the distribution of unobservables

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. 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. 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 this assumption may be unrealistic in

[18] 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.

many applications, it is adopted frequently because it is helpful for identification for reasons that will be discussed shortly.

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

Ig $\{U_1, U_0\} \perp\!\!\!\perp V|\ X.$

Imbens and Angrist (1994) showed it is possible to identify an interpretable parameter, the LATE, even if 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 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 $\mathrm{COV}(U_1 - U_0, V)|X \neq 0.$

Figure 5 helps to illustrate the implications of Ig and EH. The top panel in the figure shows that average treatment effects are allowed to vary across observable characteristics. Ig 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 Ig. 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 given x^* and $U_D = F_V(V)$, the marginal treatment effects (MTEs), are thus constant for all U_D given x^*.

The bottom panel of Fig. 5 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 U_D. 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 treatment in the absence of the program. Finally, Fig. 6 shows that while Ig 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 Ig versus EH is shown clearly in Fig. 5. Under Ig 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 Ig 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 Ig assumption. This is not necessarily the case, though, as adding a valid instrument Z to the latent index in Eq. 5 can make Ig a more plausible assumption.

In contrast to Ig, 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.

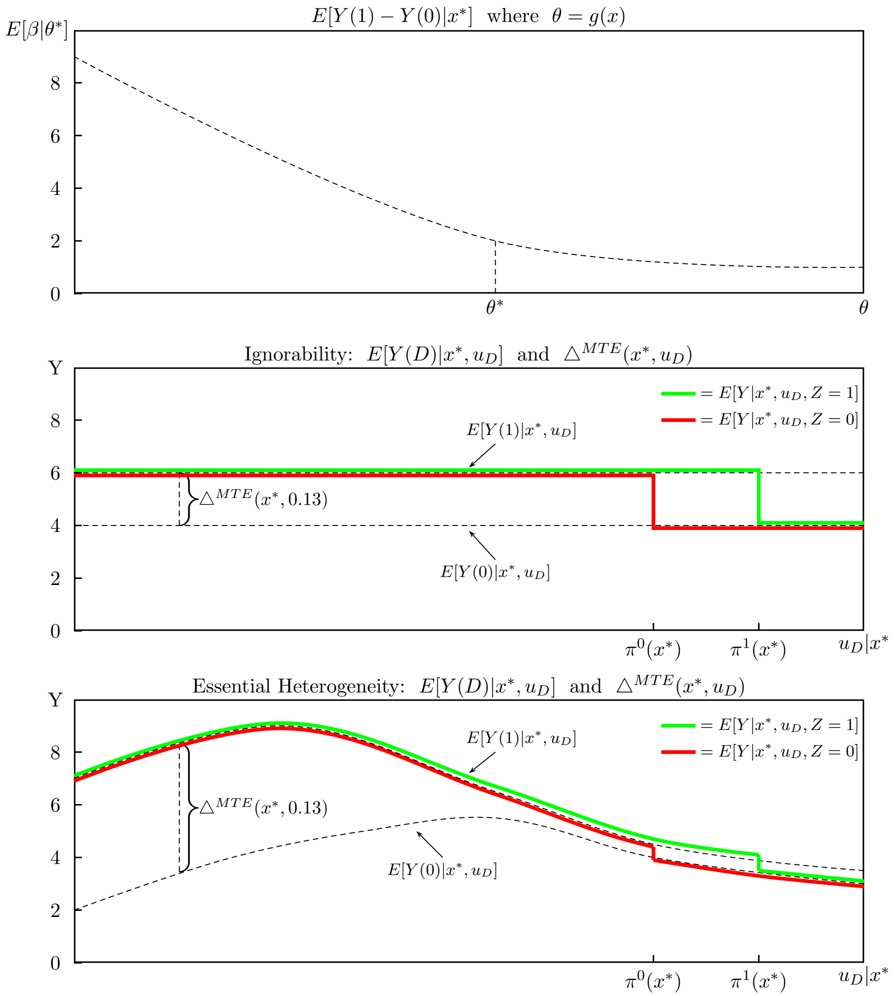

Fig. 5 Examples of Ignorability and Essential Heterogeneity

Appendix 4: External validity

Although external validity is the motivation for studying causal effects, and there is no clear reason for prioritizing internal validity over external validity (Manski 2013b), the literature has focused most formal attention on internal validity (Aliprantis 2015a). The text has adopted these priorities for the sake of publication, but here we also consider why estimated parameters will not be experiment invariant unless an assumption also holds that restricts the permissible types of peer effects (Sobel 2006). Interested readers are also directed to the careful discussions of these issues in Sobel (2006) and Ludwig et al. (2008).

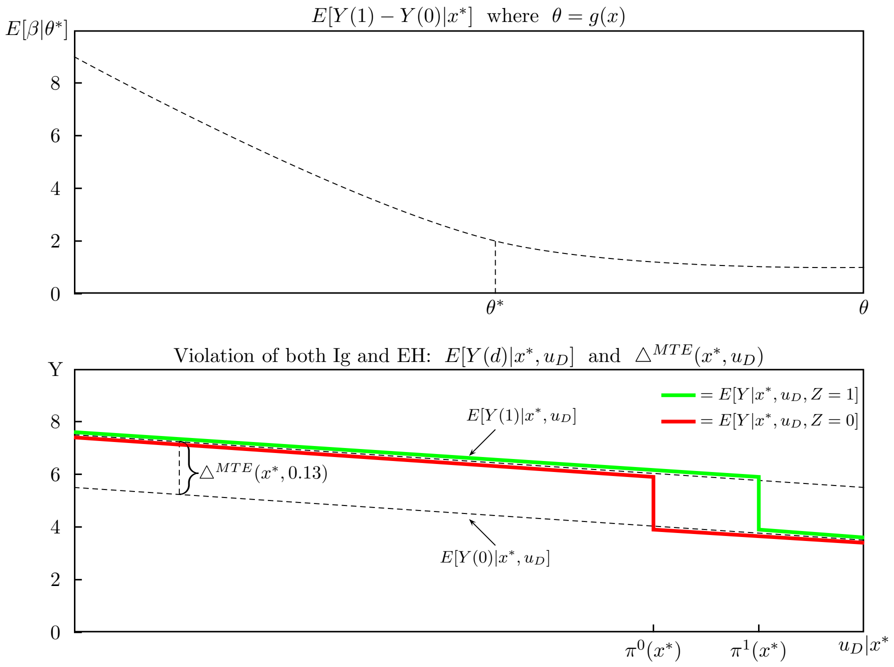

Fig. 6 Example violating both Ignorability and Essential Heterogeneity

Assumptions across and within individuals

The parameters in Sect. 4.1 are all defined conditional on the joint distribution (U, V) where we define $U \equiv (U_0, U_1)$. 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.[19] 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 $Z = 0$, here again abstracting from the Section 8 group for the sake of exposition. 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 Sect. 4.1 are defined assuming some distribution of (U, V), this implies that these parameters

[19] See p. 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.

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 Sect. 4.1 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. Under SUTVA, Ig and EH are primarily assumptions within individuals. In this case, unobservables are primarily thought to represent individual-level causal variables. Although (U, V) can represent social interactions under SUTVA, these social interactions cannot be related to treatment or assigned treatment.[20] When SUTVA is relaxed, however, Ig and EH become assumptions not only about individual-level causal variables, but also about social interactions.

A less restrictive assumption on peer effects that still keeps the effects in Sect. 4.1 identical across realizations of the randomization is that the distribution of peer effects will be identical under all realizations r. I label this as the stable peer effects assumption (SPEA):

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

Note that neither SUTVA nor SPEA is necessary to define and estimate the parameters in Sect. 4.1. However, the model illustrates how the lack of such an 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, both in terms of whether the parameters in the model are invariant to the realization of randomized voucher assignment (Sobel 2006) and in terms of parameter invariance to classes of policy interventions. Importantly, this discussion illustrates that, just like Ig or EH, parameter invariance is an assumption about the unobserved variables in the model.

Appendix 5: List of assumptions

Given the joint model of potential and outcomes and selection into treatment:

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

with

$$D = j \quad \text{if } D^* \in (C_{j-1}, C_j],$$

[20] Although this model of neighborhood effects has additional mechanisms relative to those typically included in models of social interaction, such models are still useful to consider in this context. For example, Manski (1993) and Brock and Durlauf (2007) specify models relaxing SUTVA (a) and Manski (2013a) specifies a model relaxing SUTVA (b).

there are several assumptions about the model considered throughout the paper. I list them here for the reader's reference:

A1 $\gamma_i = \gamma$ for all i and $\gamma \neq 0$
A2 $\{U_j, V\} \perp\!\!\!\perp Z \mid X$
A3 The distribution of V is absolutely continuous
A4 $E[|Y(j)||X] < \infty$ for all j
A5 $0 < Pr(D = j|X) < 1$ for all X, j
A6 $X = X_j = X_k$ almost everywhere for all $j \neq k$

D1 Treatment is moving with the aid of the program (i.e., using an MTO voucher).
D2 Treatment is moving to a high-quality neighborhood.

M1 $D_i \equiv \mathbf{1}\{\text{individual } i \text{ lives in a high-quality neighborhood}\}$
M2 $Z_i \equiv \mathbf{1}\{\text{individual } i \text{ received an MTO voucher}\}$
M3 $Y_i \equiv \mathbf{1}\{\text{individual } i \text{ is employed}\}$

NQB Neighborhood quality D is a binary function of a latent index of neighborhood quality q: $D \equiv \mathbf{1}\{q \geq q^*\}$
NQJ Neighborhood quality D is a multi-valued function of a latent index of neighborhood quality q: $D \equiv j \times \mathbf{1}\{C_{j-1} < q \leq C_j\}$
NQP Neighborhood quality q is a one-dimensional vector that is a scalar function of neighborhood poverty p: $q = \alpha p$
NQK Neighborhood quality q is a one-dimensional vector that is a linear combination of K observable neighborhood characteristics: $q = \alpha_1 X_1 + \cdots + \alpha_K X_K$

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 $j \neq i$

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

References

Aliprantis D (2015a) Covariates and causal effects: the problem of context. Federal Reserve Bank of Cleveland Working Paper 13-10R

Aliprantis D (2015b) A distinction between causal effects in structural and rubin causal models. Federal Reserve Bank of Cleveland Working Paper 15-05

Aliprantis D, Kolliner D (2015) Neighborhood poverty and neighborhood quality in the Moving to Opportunity experiment. Federal Reserve Bank of Cleveland Economic Commentary Number 2015-04. https://clevelandfed.org/~/media/content/newsroom%20and%20events/publications/economic%20commentary/2015/ec%20201504%20neighborhood%20poverty/ec%20201504%20neighborhood%20poverty%20and%20quality%20in%20the%20moving%20to%20opportunity%20experiment%20pdf.pdf?la=en

Aliprantis D, Richter FG-C (2016) Evidence of neighborhood effects from Moving to Opportunity: LATEs of neighborhood quality. Federal Reserve Bank of Cleveland Working Paper no. 12-08R. http://dionissialiprantis.com/pdfs/LATEs_of_nbd_quality_REStat1.pdf

Angrist JD (2014) The perils of peer effects. Labour Econ. doi:10.1016/j.labeco.2014.05.008

Angrist JD, Imbens GW (1995) Two-stage least squares estimation of average causal effects in models with variable treatment intensity. J Am Stat Assoc 90(430):431–442

Angrist JD, Pischke J-S (2010) The credibility revolution in empirical economics: how better research design is taking the con out of econometrics. J Econ Perspect 24(2):3–30

Brock W, Durlauf S (2007) Identification of binary choice models with social interactions. J Econom 140(1):52–75
Carneiro P, Heckman JJ, Vytlacil EJ (2011) Estimating marginal returns to education. Am Econ Rev 101(6):2754–2781
Clampet-Lundquist S, Massey DS (2008) Neighborhood effects on economic self-sufficiency: a reconsideration of the Moving to Opportunity experiment. Am J Sociol 114(1):107–143
de Souza Briggs X, Popkin SJ, Goering J (2010) Moving to Opportunity: the story of an American experiment to fight ghetto poverty. Oxford University Press, Oxford
DeLuca S, Rosenbaum JE (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. Hous Policy Debate 14(3):305–345
Friedman M (1955) The role of government in education. In: Solo R (ed) Economics and the public interest. Rutgers University Press, New Brunswick
Gennetian LA, Sciandra M, Sanbonmatsu L, Ludwig J, Katz LF, Duncan GJ, Kling JR, Kessler RC (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. Econ Policy Rev 9(2):113–140
Heckman JJ (2010) Building bridges between structural and program evaluation approaches to evaluating policy. J Econ Lit 48(2):356–398
Heckman JJ, Urzúa S, Vytlacil E (2006) Understanding instrumental variables in models with essential heterogeneity. Rev Econ Stat 88(3):389–432
Heckman JJ, Vytlacil E (2005) Structural equations, treatment effects, and econometric policy evaluation. Econometrica 73(3):669–738
Imbens G, Rubin D (2015) Causal inference for statistics, social, and biomedical sciences: an introduction. Cambridge University Press, Cambridge
Imbens GW, Angrist JD (1994) Identification and estimation of local average treatment effects. Econometrica 62(2):467–475
Keels M, Duncan GJ, Deluca S, Mendenhall R, Rosenbaum J (2005) Fifteen years later: can residential mobility programs provide a long-term escape from neighborhood segregation, crime, and poverty? Demography 42(1):51–73
Kling JR, Liebman JB, Katz LF (2007a) Experimental analysis of neighborhood effects. Econometrica 75(1):83–119
Kling JR, Liebman JB, Katz LF (2007b) Supplement to "Experimental analysis of neighborhood effects": web appendix. Econometrica 75(1):83–119
Kling JR, Ludwig J, Katz LF (2005) Neighborhood effects on crime for female and male youths: evidence from a randomized housing voucher experiment. Q J Econ 120(1):87–130
Ludwig J, Duncan GJ, Gennetian LA, Katz LF, Kessler RC, Kling JR, Sanbonmatsu L (2013) Long-term neighborhood effects on low-income families: evidence from Moving to Opportunity. Am Econ Rev 103(3):226–231
Ludwig J, Duncan GJ, Hirschfield P (2001) Urban poverty and juvenile crime: evidence from a randomized housing-mobility experiment. Q J Econ 116(2):655–679
Ludwig J, Kling JR (2007) Is crime contagious? J Law Econ 50(3):491–518
Ludwig J, Liebman JB, Kling JR, Duncan GJ, Katz LF, Kessler RC, Sanbonmatsu L (2008) What can we learn about neighborhood effects from the Moving to Opportunity experiment? Am J Sociol 114(1):144–188
Manski CF (1993) Identification of endogenous social effects: the reflection problem. Rev Econ Stud 60(3):531–542
Manski CF (2011) Choosing treatment policies under ambiguity. Annu Rev Econ 3:25–49
Manski CF (2013a) Identification of treatment response with social interactions. Econom J 16(1):S1–S23
Manski CF (2013b) Public policy in an uncertain world: analysis and decisions. Harvard University Press, Cambridge
Mendenhall R, DeLuca S, Duncan G (2006) Neighborhood resources, racial segregation, and economic mobility: results from the Gautreaux program. Soc Sci Res 35(4):892–923
Minnesota Population Center (2004) National historical geographic information system (pre-release version 0.1 ed.). University of Minnesota, Minneapolis. http://www.nhgis.org

Orr LL, Feins JD, Jacob R, Beecroft E, Sanbonmatsu L, Katz LF, Liebman JB, Kling JR (2003) Moving to Opportunity: interim impacts evaluation. US Department of Housing and Urban Development, Office of Policy Development and Research, Washington, DC

Pearl J (2009) Causality: models, reasoning and inference, 2nd edn. Cambridge University Press, Cambridge

Pinto R (2014) Selection bias in a controlled experiment: the case of Moving to Opportunity. University of Chicago, Chicago

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

Quigley JM, Raphael S (2008) Neighborhoods, economic self-sufficiency, and the MTO program. Brook Whart Pap Urban Aff 8(1):1–46

Rosenbaum JE (1995) Changing the geography of opportunity by expanding residential choice: lessons from the Gautreaux program. Hous Policy Debate 6(1):231–269

Rubin DB (1978) Bayesian inference for causal effects: the role of randomization. Ann Stat 6(1):34–58

Rubinowitz LS, Rosenbaum JE (2000) Crossing the class and color lines: from public housing to white suburbia. University of Chicago Press, Chicago

Sampson RJ (2008) Moving to inequality: neighborhood effects and experiments meet social structure. Am J Sociol 114(1):189–231

Sampson RJ (2012) Great American city: Chicago and the enduring neighborhood effect. The University of Chicago Press, Chicago

Sampson RJ, Sharkey P, Raudenbush SW (2008) Durable effects of concentrated disadvantage on verbal ability among African-American children. Proc Natl Acad Sci USA 105(3):845–852

Sanbonmatsu L, Kling JR, Duncan GJ, Brooks-Gunn J (2006) Neighborhoods and academic achievement: results from the Moving to Opportunity experiment. J Hum Resour 41(4):649–691

Sanbonmatsu L, Marvakov J, Potter NA, Yang F, Adam E, Congdon WJ, Duncan GJ, Gennetian LA, Katz LF, Kling JR, Kessler RC, Lindau ST, Ludwig J, McDade TW (2012) The long-term effects of moving to opportunity on adult health and economic self-sufficiency. Cityscape 14(2):109–136

Sobel ME (2006) What do randomized studies of housing mobility demonstrate? Causal inference in the face of interference. J Am Stat Assoc 101(476):1398–1407

Votruba ME, Kling JR (2009) Effects of neighborhood characteristics on the mortality of black male youth: evidence from Gautreaux, Chicago. Soc Sci Med 68(5):814–823

Wilson WJ (1987) The truly disadvantaged: the inner city, the underclass, and public policy. University of Chicago, Chicago