

Evidence of Neighborhood Effects from Moving to Opportunity: LATEs of Neighborhood Quality

Dionissi Aliprantis Francisca G.-C. Richter*
Federal Reserve Bank of Cleveland Case Western Reserve University

First Version: March 2012. This Version: May 5, 2016

Abstract: This paper finds evidence of positive neighborhood effects on adult labor market outcomes using the Moving to Opportunity (MTO) housing mobility experiment. Our results stand in such sharp contrast to the current literature because our analysis focuses on outcomes of the subpopulation induced by the program to move to a higher quality neighborhood, while previous analyses have focused on outcomes of either the entire population or the subpopulation induced by the program to move. We propose and implement a new strategy for identifying heterogeneous transition-specific effects that exploits the identification of the idiosyncratic, unobserved component of a neighborhood choice model. We estimate Local Average Treatment Effects (LATEs) of the change in quality most commonly induced by MTO vouchers, between the first and second deciles of the national distribution of neighborhood quality. Although MTO vouchers induced much larger changes in neighborhood quality than standard Section 8 vouchers, these LATEs only pertain to a subpopulation representing about nine percent of program participants.

Keywords: Neighborhood Effect, Local Average Treatment Effect, Moving to Opportunity
JEL Classification Numbers: C31, C36, C50, D04, I38, R23

*Address: Federal Reserve Bank of Cleveland, Research Department, PO Box 6387, Cleveland, OH 44101-1387, USA. Phone: +1(216)579-3021. E-mail: *dionissi.aliprantis@clev.frb.org*, *francisca.richter@case.edu*.

Acknowledgements: The authors thank Jeffrey Kling, Joel Elvery, Fabian Lange, Jon James, Ben Craig, Nathaniel Baum-Snow, Karim Chalak, Stephen Ross, Kyle Hood, Carlos Lamarche, Alvin Murphy, and several anonymous referees for helpful comments, as well as seminar participants at GATE/Lyon II, the Kentucky-Cleveland Fed Economics of Education Workshop, the 2013 Midwest Economics Association conference, and the 2013 Urban Economics Association conference. We also thank Emily Burgen and Nelson Oliver for valuable research assistance, and Paul Joice at HUD for his assistance with the data set. The views stated herein are those of the authors and are not necessarily those of the Federal Reserve Bank of Cleveland or the Board of Governors of the Federal Reserve System.

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. One prominent theory is that these differences in outcomes can be explained by effects from living in a poor, segregated, and socially isolated neighborhood (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.

Because households endogenously sort into neighborhoods, researchers have sought to identify neighborhood effects using the exogenous variation in neighborhoods induced by housing mobility programs. One of the best known housing mobility programs is the Gautreaux program, which was designed to desegregate public housing in Chicago and relocated public housing residents in a quasi-random manner using housing vouchers. 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)).

The Moving to Opportunity (MTO) housing mobility experiment sought to replicate the quasi-experimental results from Gautreaux with a randomized experiment. MTO gave households living in high-poverty neighborhoods in five US cities the ability to enter a lottery for housing vouchers to be used in low-poverty neighborhoods. In a tremendous disappointment, effects from MTO were much smaller than effects from Gautreaux.

The lack of effects from MTO has been interpreted as evidence against the theory that neighborhood characteristics influence individuals’ outcomes. For example, Ludwig et al. (2013) interpret the lack of effects from MTO as evidence “Contrary to the widespread view that living in a disadvantaged inner-city neighborhood depresses labor market outcomes” (p 288). This interpretation links effects from MTO and effects from neighborhoods under the assumption that MTO induced changes in neighborhood characteristics large enough 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 on behavior” (Ludwig et al. (2008), pp 181-182). Aliprantis (2015a) explicitly states the assumptions necessary to make this link between effects from MTO and effects from neighborhood environments, and presents empirical evidence that these assumptions are unlikely to hold in the case of MTO.¹

Not only did MTO induce a select subset of households to move to better neighborhoods, but improvements in neighborhood quality were small even for these movers. While this helps to explain the small effects from the program, it limits the usefulness of ITT and TOT effects for inferring the existence of neighborhood effects. Rather than attempting to learn about neighborhood effects from ITT and TOT effects of the MTO program, this paper identifies Local Average Treatment Effects

¹Related analyses can be found in Clampet-Lundquist and Massey (2008), Sampson (2008), Quigley and Raphael (2008), and Pinto (2014).

(LATEs) of neighborhood quality on adult outcomes using the random variation in quality induced by the MTO program together with an explicit model of households' selection into neighborhood quality.² We find that moving to a higher quality neighborhood had large, positive effects on employment, labor force participation, household income, and welfare receipt. Although effects on Body Mass Index (BMI) and individual earnings are not statistically significant, these effects are also estimated to be large and positive. We find no evidence from MTO against the theory that increasing neighborhood quality improves adult outcomes.

We find evidence of neighborhood effects from MTO by focusing our analysis on the subpopulation induced by MTO to move to higher quality neighborhoods. Earlier studies have failed to find these effects because they have focused on the outcomes for broader populations and used more restrictive models of neighborhood effects. For example, estimates of neighborhood effects on adult labor market outcomes are absent from the most prominent analysis of MTO because the program was found to have little effect on such outcomes (Kling et al. (2007)).³ However, our focus on moves to higher quality illustrates just how few parameters can be identified with the MTO data. Not only do our LATE estimates pertain to about nine percent of program participants, but relative to the national distribution, our LATEs of neighborhood quality only pertain to moves between low quality neighborhoods.

Thus while our results provide support for the idea that efforts to deconcentrate poverty are well-grounded, our estimates' lack of generality is itself evidence that policies ought to be carefully designed to achieve policy-makers' objectives. Despite the fact that households were more likely to move with Section 8 vouchers than MTO vouchers, changes in neighborhood quality were much smaller for Section 8 movers than for MTO movers, and variation by site was large. Since only about a quarter of eligible households are currently able to obtain a Section 8 housing voucher (Sard and Fischer (2012)), an area for future research is understanding what types of restrictions on voucher use might optimize the extent to which households are able to realize positive neighborhood effects through the subsidy, and which of these restrictions might be feasible to implement (McClure (2010)).

A final caveat when interpreting the policy implications of our results is that our estimated LATEs are total effects, encompassing the direct effects from changes in neighborhoods and from other programs. Factors other than MTO influencing neighborhood decline or revitalization in this time period are not explicitly modeled in our analysis. Nor are other programs that might differentially impact the voucher and control groups in the program (Heckman et al. (2000)), such as job training programs, education reform, other housing programs like HOPE VI, or welfare reform. The difficulty of separately identifying the direct effect of neighborhood environments points to the methodological limitations of conducting randomized, but not controlled, experiments in social settings (Deaton (2010), Aliprantis (2015b)).

²We leave the analysis of child and youth outcomes for future work because we believe schools are likely to be the most important "neighborhood" for children (Oreopoulos (2003)), and there is difficulty in measuring school quality in the MTO data.

³Kling et al. (2007) focus their attention on "outcomes that exhibit significant treatment effects" (p. 83).

We begin our analysis by proposing a new strategy for identifying transition-specific LATEs with the aid of an ordered choice model in the absence of transition-specific instruments. Alternative identification strategies yield parameters that are weighted averages of such effects (Angrist and Imbens (1995)), or else require transition-specific instruments for each margin of choice (Heckman et al. (2006)). Our identification strategy is positioned between these two extremes, in that it is both economically interesting and empirically feasible. The key insight of the identification strategy is that by observing the continuous neighborhood quality associated with each program participant, it is possible to identify the unobserved, idiosyncratic component of their latent index in the neighborhood choice model.⁴

Our model specification and empirical implementation is guided by the desire to build on the analysis in Kling et al. (2007) in three important ways. While the notation can be burdensome at times, we stress that this does not imply stronger assumptions than those imposed on the data by more concisely described models.

First and foremost, we aim to focus on outcomes of the subgroup of households that were actually induced by the program to move to higher quality neighborhoods. With this goal, the model must be able to accurately characterize selection into neighborhoods of various quality levels. An important finding from our estimated ordered choice model of selection into neighborhood quality is that although households were more likely to move with Section 8 vouchers than MTO vouchers, variation by site was large, and changes in neighborhood quality were much smaller for Section 8 movers than for MTO movers in all sites except Boston (studied extensively in Galiani et al. (2015)).

Second, the model is also motivated by a desire to relax the assumption that effects are homogeneous conditional on observed characteristics. Simultaneously achieving the first two objectives of our analysis requires that our model specification balance the asymmetry inherent to Instrumental Variable (IV) identification strategies, which allow for general heterogeneity in response to treatment while restricting the patterns of heterogeneity in response to the instrument from complete generality (Heckman et al. (2006), Imbens and Angrist (1994), Aliprantis (2012)). We do this by modeling heterogeneity in response to the instrument with a finite mixture model satisfying a monotonicity assumption.

Third, the empirical implementation of the model allows us to relax assumptions about the precise neighborhood characteristics that influence outcomes. We create and utilize a linear index of neighborhood quality informative about several neighborhood characteristics in addition to the neighborhood poverty rate. Using this precise quality index makes assumptions about the neighborhood characteristics that affect outcomes. Unfortunately, doing so is simply unavoidable: The literature using neighborhood poverty rate as the index measuring quality also makes assumptions

⁴The case of a binary treatment is originally developed in Imbens and Angrist (1994), Heckman and Vytlačil (1999), and Heckman and Vytlačil (2005). Identification strategies related to ours are developed in Heckman et al. (2006), Lochner and Moretti (2015), Nekipelov (2011), and Kim and Petrin (2013). Identification strategies creating comparable groups based on choice-revealed unobservables have been implemented in applied work like Altonji and Mansfield (2014), Fletcher and Ross (2012), Fu and Ross (2013), Bayer and Ross (2009), and Dale and Krueger (2002).

about the neighborhood characteristics that affect outcomes. We believe that our index of quality imposes assumptions less likely to be violated than the index of neighborhood poverty (A lengthy discussion can be found in Aliprantis (2015a)).

The paper proceeds as follows: Section 2 specifies our joint model of neighborhood choice and potential outcomes, defining the treatment effect parameters we seek to identify in terms of this model. Section 3 presents our strategy for identifying these effects, with Appendix A presenting a formal justification, and other sections of the Appendix offering intuition and interpretation of our model. Section 4 describes the MTO housing mobility program, the data used in estimation, and descriptive statistics of those data. Section 5 presents our empirical specification and estimation results, with Sections 5.1 and 5.2 focused on the ordered choice model, and Section 5.3 focused on neighborhood effects. Section 6 discusses our interpretation of the results, and Section 7 concludes.

2 Model

2.1 A Joint Model of Neighborhood Choice and Potential Outcomes

We adopt a continuous model of neighborhood choice. At baseline, MTO participants tended to live in some of the lowest quality neighborhoods in the US. As a result, we view neighborhood selection in terms of households starting at some very low level of quality $q = \delta > 0$, and then choosing to move to a higher quality neighborhood as long as the marginal benefit of doing so is positive. We assume there is a latent index defining the marginal net benefit of moving to a higher quality neighborhood, where the benefit is a function of voucher assignment and take-up, metro area, and household characteristics:

$$\mathbf{A1:} \quad MB_i(q) = \mu(X_i) - C(q) - V_i + \gamma^M(q)Z_i^M\tau_i^M \quad \text{where}$$

X_i is a vector of household characteristics at baseline;

$C(q)$ is a continuous and increasing function of $q \in [\delta, 50]$ representing marginal costs due to specific metro characteristics;

V_i accounts for the household's unobserved characteristics that influence marginal cost such as transportation and childcare needs;

Z_i^M is an indicator of voucher assignment;

$\tau_i^M = \mathbf{1}\{\mu^M(X_i) - V_i^M \geq 0\}$ is an indicator of voucher take-up;

V_i^M accounts for the household's unobserved characteristics that influence voucher take-up;

$\gamma^M(q)$ is a continuous function of $q \in [\delta, 50]$ representing marginal benefits specific to voucher use.

We model potential outcomes in terms of discrete levels of neighborhood quality,

$$D = \begin{cases} 1 & \text{if } q \in [\underline{q}_1^D, \bar{q}_1^D); \\ \vdots & \vdots \\ J & \text{if } q \in [\underline{q}_J^D, \bar{q}_J^D), \end{cases} \quad (1)$$

where potential outcomes are

$$\mathbf{A2:} \quad Y_{ij} = \mu_j(X_i) + U_{ij} \quad \text{for } j = 1, \dots, J.$$

We add the following assumptions to A1-A2 that are similar to those found in the ordered choice model developed in Heckman et al. (2006):

$$\mathbf{A3:} \quad (X_i, V_i, V_i^M, U_{ij}) \perp\!\!\!\perp \mathbf{Z}_i \text{ for all } j = 1, \dots, J$$

$$\mathbf{A4:} \quad \mathbb{E}[|Y_j|] < \infty \text{ for all } j = 1, \dots, J$$

$$\mathbf{A5:} \quad \gamma^M(q) > 0 \text{ for all } q \in [\delta, 50]$$

We emphasize that no assumptions are made about potential outcomes through the U_{ij} other than the valid instrument assumption (A3) and the assumption of integrability (A4).

Our analysis will focus on the j to $j + 1$ transition-specific Local Average Treatment Effect (LATE) for the experimental MTO voucher:

$$\Delta_{j,j+1}^{LATE}(Z^M) \equiv \mathbb{E}[Y_{j+1} - Y_j \mid D(Z^M = 0) = j, D(Z^M = 1) = j + 1].$$

We note that this parameter is instrument-specific, maintaining this distinguishing feature of the binary LATE parameter.⁵

2.2 Interpretation and Discussion of Assumptions

We think of our model as an “as if” abstraction, where the marginal benefit function summarizes the factors determining neighborhood of residence by the time of the interim survey. Note that since the interim survey was conducted four to eight years after randomization, this left sufficient time for households to have moved more than once. In this sense, the marginal benefit function can be thought to capture all of the factors leading to the selection of the neighborhood of residence at the time of the interim survey.

Many of the abstractions in our model can be motivated empirically, or arise due to measurement issues. For example, our modeling of neighborhood selection as beginning at a very low level of quality can be motivated empirically, since the median neighborhood quality at baseline for MTO participants was below the first percentile of the national distribution (Figure 4b). Another example

⁵The fact that the LATE is instrument-specific has received much discussion in the literature; see Imbens and Angrist (1994), Heckman and Urzúa (2010), Heckman and Vytlacil (2005).

is the $C(q)$, which may include increased search costs of housing in higher quality, higher priced neighborhoods due to a reduced supply of units available at the metro-level Fair Market Rent (FMR).⁶ We do not explicitly include rents or house prices in our model because households pay 30 percent of their income regardless of the unit’s rent, as long as it is a Section 8 eligible unit. Thus the probability of finding a Section 8 eligible unit at the metro-level FMR is the key to the household’s choice. This probability is difficult to measure; Appendix E presents further discussion.

The marginal benefit function allows for idiosyncratic, heterogeneous response to the instrument. The contribution of moving with a voucher $\gamma_i^M(q)Z_i^M\tau_i^M$ can be interpreted as a reduction in the cost of improving quality specifically occurring at quality level q brought about by offer ($Z_i = 1$) and take-up ($\tau_i = 1$) of the voucher. Households that receive a voucher ($Z_i = 1$) but do not move with it ($\tau_i = 0$) are those for which the voucher does not represent a cost reduction of moving.

Although we would ideally be able to allow for even more general heterogeneity in response to the instrument, the bivariate latent classes in A1 relax the assumption of homogeneous responses to vouchers, and are a tractable, interpretable way of modeling this heterogeneity (See Greene et al. (2008) for a related discussion.). Furthermore, while the type of heterogeneity allowed under A1 is not completely general, comparable restrictions are necessary for Instrumental Variables (IV) to identify causal parameters of interest, as are additional modeling assumptions (Vytlacil (2002), Vytlacil (2006)). While IV identification strategies can allow for general heterogeneity of outcomes, the patterns of heterogeneity are not completely general in response to the instrument under which IV identifies causal effects (Heckman et al. (2006), Imbens and Angrist (1994), Aliprantis (2012)).

A5 is a monotonicity assumption, where the existence of some (X_i, V_i^M) such that $\tau_i^M = 1$ ensures that the instrument is relevant. The strength of the instrument is determined both by how many households move with the voucher $\tau_i^M = 1$, as well as the function $\gamma^M(q)$ describing the cost reduction represented by the voucher.

The combined benefits and costs of moving to neighborhoods of various levels of quality are determined by social interactions, market forces, and political processes. Since the number of program movers is small relative to neighborhood size, our model assumes that movers do not influence the market price of neighborhood quality, the technology by which resources are produced, or the political process by which resources are distributed. These partial equilibrium assumptions are not only reasonable, but also greatly facilitate estimation relative to a general equilibrium model of neighborhood effects with social interactions, market forces, and political processes. Consider that we are able to estimate our joint model of both selection and outcomes, when it is difficult to estimate models of selection (Manski (1993), Brock and Durlauf (2007)) or outcomes (Manski (2013a)) even when modeled separately and focused only on the mechanism of social interaction effects.

We interpret effects of neighborhood quality as the result of social interactions and

⁶The FMR is usually set at 40th percentile of local gross rents for typical non-substandard rental units, adjusted by household size.

neighborhood-level resources (by way of the technology utilizing and distributing resources across and within neighborhoods). In our model, each level of neighborhood quality is thought to represent, on average, homogeneous types of social interactions and resources available at the neighborhood level. We assume agents experience neighborhood quality and cannot determine it because we empirically define neighborhoods as census tracts, which contain about 4,000 residents on average. If we were thinking about smaller reference groups like social groups in classrooms, we would be more interested in incorporating the endogenous formation of reference groups into the model. Since we are dealing with large N , we believe our partial equilibrium abstraction is appropriate.

We discuss the invariance of the estimated LATE parameter in Section 6. One related point about invariance is especially important in our application. If the Stable Unit Treatment Value Assumption (SUTVA) fails to hold, say due to social interactions, then the joint distribution of $(V, V^S, V^M, U_1, \dots, U_J)$, and therefore the LATE, need not be invariant to different realizations of the same policy, even when randomized (Sobel (2006), Aliprantis (2015a)).

3 Identification

3.1 First Stage: Identifying the Neighborhood Choice Model

This Section discusses how we identify the continuous neighborhood choice model using an ordered choice model, with Appendix A providing a formal justification.⁷ Intuitively, the reason we might be interested in such an identification strategy is that the cost function $C(q)$ will assume a different curvature for any given parametric distribution of V_i , and estimating a ordered choice model allows for a flexible estimation of the curvature of this continuous cost function.⁸ As we will see in Section 3.2, in the absence of transition-specific instruments, identifying the V_i from the neighborhood choice model is essential to our approach to identifying transition-specific LATEs.

For the sake of exposition, we first consider identification of the continuous model parameters for the control group, and then expand on this for the experimental MTO group. Define a new discretization of the continuous measure of quality as:

$$Q_i = \begin{cases} 1 & \text{if } q_i \in [\underline{q}_1^Q, \bar{q}_1^Q); \\ \vdots & \quad \quad \quad \vdots \\ K & \text{if } q_i \in [\underline{q}_K^Q, \bar{q}_K^Q), \end{cases} \quad (2)$$

and a corresponding discretization of $MB_i(q)$ characterizing the marginal benefit of moving from

⁷Specifically, we prove that the expected value of the \widehat{V}_i estimates from Equation 12 constructed from an ordered choice model as detailed in this Section can be made arbitrarily close to the V_i from the continuous model in A1 as the discretization of quality defined in Equation 2 is defined in terms of finer and finer partitions.

⁸Any chosen parametric distribution of V_i will serve as a normalization, and the other features of the model, including the cost function, will all be parameterized relative to the chosen distribution. For example, assuming that V_i is distributed according to a standard normal, a normal with mean zero and variance of 2, or a logistic distribution with given scale and location parameters will all result in different $C(q)$ functions.

the discrete quality level k to $k + 1$:

$$MB_{ik} = MB_i(\bar{q}_k^Q) = \mu(X_i) - C(\bar{q}_k^Q) - V_i.$$

The optimal quality level $k^* \in \{1, \dots, K\}$ satisfies the following condition

$$Q_i = k^* \iff MB_{ik^*} < 0 \leq MB_{ik^*-1}, \quad (3)$$

which under our specification is equivalent to

$$Q_i = k^* \iff \mu(X_i) - C(\bar{q}_{k^*}^Q) < V_i \leq \mu(X_i) - C(\bar{q}_{k^*-1}^Q). \quad (4)$$

Note that Conditions 3 and 4 are just the standard ordered choice model condition

$$Q_i = k^* \iff \mu(X_i) - C_{k^*} < V_i \leq \mu(X_i) - C_{k^*-1} \quad (5)$$

where C_k is the continuous function $C(q)$ evaluated at the threshold \bar{q}_k^Q , or where $C_k = C(\bar{q}_k^Q)$.

From Condition 5, we know that for the control group,

$$P(Q_i = k | X_i) = F_V(\mu(X_i) - C_{k-1}) - F_V(\mu(X_i) - C_k),$$

and so maximum likelihood estimation of an ordered choice model yields estimates of $\hat{\mu}(X_i)$ and $\{\hat{C}_k\}_{k=1}^{K-1}$. Linear interpolation between the \hat{C}_k and \hat{C}_{k+1} allows us to construct an estimate of the continuous cost function:

$$\hat{C}(q) = \hat{C}_k + (q - \bar{q}_k^Q) \left(\frac{\hat{C}_{k+1} - \hat{C}_k}{\bar{q}_{k+1}^Q - \bar{q}_k^Q} \right) \quad \text{for } q \in (\bar{q}_k^Q, \bar{q}_{k+1}^Q). \quad (6)$$

This procedure, based on the fact that $C_k = C(\bar{q}_k^Q)$, is the key to obtaining estimates of the parameters of the continuous model through the estimation of a discrete ordered choice model. As the discretization of quality defined in Equation 2 is defined in terms of finer and finer partitions, this approximation to the continuous model converges to the continuous model (Appendix A).

Figure 1 below previews our empirical results by showing the continuous cost function generated by linear interpolation between the cut points of an ordered choice model, and also illustrates that the discretization of quality used to estimate the neighborhood choice model (Equation 2) need not be the same as the discretization used for potential outcomes (Equation 1).

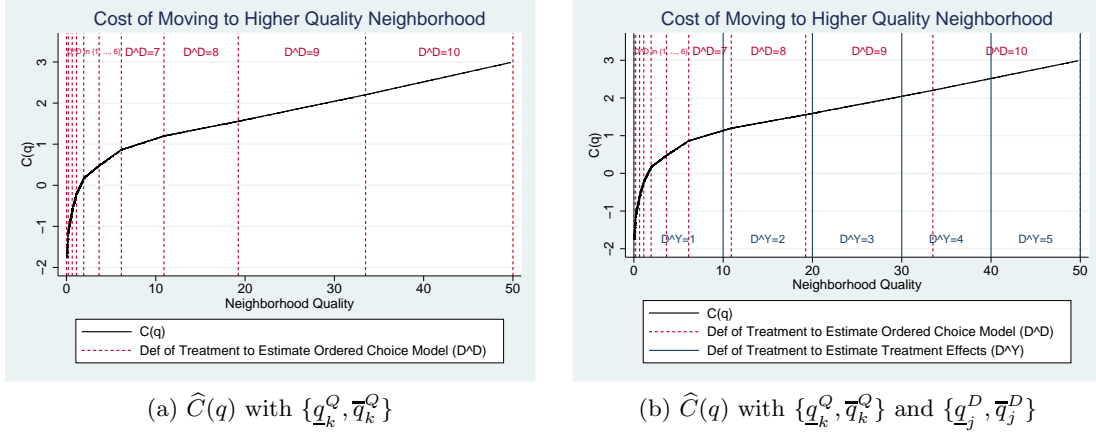


Figure 1: Cost Function

Note: The dashed red vertical lines in (a) show the discretization of quality used to define Q_i , and the cost function $C(q)$ is constructed by linearly interpolating between the cutpoints estimated in an ordered probit model. The solid blue vertical lines in (b) show the discretization of quality used to define D_i as used in the model of potential outcomes.

Turning to the objective of identifying each household's V_i , note that utility is maximized where the decreasing marginal benefit function $MB_i(q)$ crosses zero. This First Order Condition (FOC) implies that

$$V_i = \mu(X_i) - C(q_i). \quad (7)$$

Thus, given estimates of the right hand side of Equation 7, for the control group we can identify each household's unobserved component in the ordered choice model as

$$\widehat{V}_i = \widehat{\mu}(X_i) - \widehat{C}(q_i). \quad (8)$$

For identification of the parameters pertaining to the experimental MTO voucher recipients, note that for the movers receiving the MTO voucher ($Z_i^M = 1$ and $\tau_i^M = 1$), we have the condition that

$$Q_i = k^* \iff \mu(X_i) - C_{k^*} + \gamma_{k^*}^M < V_i \leq \mu(X_i) - C_{k^*-1} + \gamma_{k^*-1}^M. \quad (9)$$

Since all of the parameters in Equation 9 except the γ_k^M are identified by the control group, the γ_k^M are then identified by the movers receiving the MTO voucher. After obtaining estimates of $\{\gamma_k^M\}_{k=1}^{K-1}$ via maximum likelihood estimation of the ordered choice model, we can again linearly interpolate to identify the continuous function

$$\widehat{\gamma}^M(q) = \widehat{\gamma}_k^M + (q - \bar{q}_k^Q) \left(\frac{\widehat{\gamma}_{k+1}^M - \widehat{\gamma}_k^M}{\bar{q}_{k+1}^Q - \bar{q}_k^Q} \right) \quad \text{for } q \in (\bar{q}_k^Q, \bar{q}_{k+1}^Q). \quad (10)$$

Note that the general FOC is

$$V_i = \mu(X_i) + \gamma^M(q_i) Z_i^M \tau_i^M - C(q_i), \quad (11)$$

and that we observe in the data whether households move ($\tau_i^M \in \{0, 1\}$) when given an MTO

voucher ($Z_i^S = 1$), and when households are not given an MTO voucher $Z_i^M = 0 \Rightarrow \gamma^M(q_i)Z_i^M\tau_i^M = 0$. Thus, all of the parameters on the righthand side of Equation 11 are identified, and thus so are the V_i :

$$\widehat{V}_i = \widehat{\mu}(X_i) - \widehat{C}(q_i) + \widehat{\gamma}^M(q)Z_i^M\tau_i^M. \quad (12)$$

3.2 Second Stage: Identifying LATEs of Neighborhood Quality

In the empirical analysis it will be convenient to refer interchangeably to parameters and sets defined in terms of V_i and U_{Di} , where $U_{Di} \equiv F_V(V_i)$. We are interested in the regions of observed and unobserved characteristics, $(\mu(X_i), U_{Di})$, corresponding to households that would select into treatment level j in the absence of a voucher, but could possibly select into treatment level $j + 1$ if given an MTO voucher. To focus on this group, we first define the identification support set as:

$$\mathcal{S}_j^M \equiv \left\{ \begin{array}{l} (\mu(X_i), U_{Di}) \quad | \quad D_i(Z_i^M = 0) = j, \quad D_i(Z_i^M = 1) \in \{j, j + 1\}, \\ \Pr(D_i = j + 1 | \mu(X_i), U_{Di}, Z_i^M = 1) > 0 \end{array} \right\} \subset \mathcal{M} \times [0, 1].$$

We stress that membership in the identification support set \mathcal{S}_j^M can only be determined after identifying each household's V_i as established in the previous Section. Determining \mathcal{S}_j^M gives us the ability to identify transition-specific LATEs with a single instrument.

When offered an MTO voucher ($Z_i^M = 1$), a household in \mathcal{S}_j^M can choose to respond in one of three mutually exclusive options, which we denote as follows:

- 1 $S_j^M(\tau_i^M = 0, D_i = j)$ for not using the voucher, remaining in a neighborhood of quality j ;
- 2 $S_j^M(\tau_i^M = 1, D_i = j)$ for moving with the voucher, but not to a higher quality neighborhood;⁹
- 3 $S_j^M(\tau_i^M = 1, D_i = j + 1)$ for moving with the voucher to a higher quality neighborhood.

We can then classify households into non-compliers and compliers as follows:

$$\mathcal{NC}_j^M = S_j^M(\tau_i^M = 0, D_i = j) \sqcup S_j^M(\tau_i^M = 1, D_i = j)$$

$$\mathcal{C}_j^M = S_j^M(\tau_i^M = 1, D_i = j + 1),$$

where \sqcup represents a disjoint union. We denote the probability of being a complier conditional on being in \mathcal{S}_j^M as $\pi(\mathcal{C}_j^M)$, and note that $\mathcal{S}_j^M = \mathcal{NC}_j^M \sqcup \mathcal{C}_j^M$.

To facilitate exposition, we adopt the notation

$$\begin{aligned} i \in \mathcal{C}_j^M & \text{ to signify } (\mu(X_i), U_{Di}, \mu^M(X_i), V_i^M) \in \mathcal{C}_j^M \\ i \in \mathcal{NC}_j^M & \text{ to signify } (\mu(X_i), U_{Di}, \mu^M(X_i), V_i^M) \in \mathcal{NC}_j^M \\ i \in \mathcal{S}_j^M & \text{ to signify } (\mu(X_i), U_{Di}) \in \mathcal{S}_j^M. \end{aligned}$$

⁹One possibility generating this case is that the household moves to a low-poverty neighborhood of the same quality. Another possibility is that because the interim study was conducted four to eight years after randomization, households could have moved more than once, with their final move being to a neighborhood of quality level j .

The Wald estimator applied to the subsample of experimental and control households in \mathcal{S}_j^M identifies the j to $j + 1$ transition-specific LATE:

$$\begin{aligned}
& \frac{\mathbb{E}[Y \mid Z^M = 1, i \in \mathcal{S}_j^M] - \mathbb{E}[Y \mid Z^M = 0, i \in \mathcal{S}_j^M]}{\mathbb{E}[D \mid Z^M = 1, i \in \mathcal{S}_j^M] - \mathbb{E}[D \mid Z^M = 0, i \in \mathcal{S}_j^M]} \tag{13} \\
&= \frac{\left\{ \mathbb{E}[Y_{j+1} \mid i \in \mathcal{C}_j^M] \pi(\mathcal{C}_j^M) + \mathbb{E}[Y_j \mid i \in \mathcal{N}\mathcal{C}_j^M] \pi(\mathcal{N}\mathcal{C}_j^M) \right\} - \left\{ \mathbb{E}[Y_j \mid i \in \mathcal{C}_j^M] \pi(\mathcal{C}_j^M) + \mathbb{E}[Y_j \mid i \in \mathcal{N}\mathcal{C}_j^M] \pi(\mathcal{N}\mathcal{C}_j^M) \right\}}{\left\{ (j+1)\pi(\mathcal{C}_j^M) + (j)\pi(\mathcal{N}\mathcal{C}_j^M) \right\} - \left\{ (j)\pi(\mathcal{C}_j^M) + (j)\pi(\mathcal{N}\mathcal{C}_j^M) \right\}} \\
&= \frac{\left(\mathbb{E}[Y_{j+1} \mid i \in \mathcal{C}_j^M] - \mathbb{E}[Y_j \mid i \in \mathcal{C}_j^M] \right) \pi(\mathcal{C}_j^M) + \left(\mathbb{E}[Y_j \mid i \in \mathcal{N}\mathcal{C}_j^M] - \mathbb{E}[Y_j \mid i \in \mathcal{N}\mathcal{C}_j^M] \right) \pi(\mathcal{N}\mathcal{C}_j^M)}{\left[(j+1) - j \right] \pi(\mathcal{C}_j^M) + \left[j - j \right] \pi(\mathcal{N}\mathcal{C}_j^M)} \\
&= \frac{\mathbb{E}[Y_{j+1} - Y_j \mid i \in \mathcal{C}_j^M] \pi(\mathcal{C}_j^M)}{\pi(\mathcal{C}_j^M)} \\
&= \mathbb{E}[Y_{j+1} - Y_j \mid i \in \mathcal{C}_j^M] \\
&\equiv \Delta_{j,j+1}^{LATE}(Z^M).
\end{aligned}$$

The colors are added above to facilitate following the algebra. To justify Equation 13, note that

$$\begin{aligned}
& \mathbb{E}[Y \mid Z^M = 1, i \in \mathcal{S}_j^M] \\
&= \mathbb{E}[Y \mid Z^M = 1, i \in \mathcal{C}_j^M, i \in \mathcal{S}_j^M] Pr(i \in \mathcal{C}_j^M \mid i \in \mathcal{S}_j^M) \\
&\quad + \mathbb{E}[Y \mid Z^M = 1, i \in \mathcal{N}\mathcal{C}_j^M, i \in \mathcal{S}_j^M] Pr(i \in \mathcal{N}\mathcal{C}_j^M \mid i \in \mathcal{S}_j^M) \tag{14} \\
&= \mathbb{E}[Y \mid Z^M = 1, i \in \mathcal{C}_j^M, i \in \mathcal{S}_j^M] \pi(\mathcal{C}_j^M) + \mathbb{E}[Y \mid Z^M = 1, i \in \mathcal{N}\mathcal{C}_j^M, i \in \mathcal{S}_j^M] \pi(\mathcal{N}\mathcal{C}_j^M) \tag{15} \\
&= \mathbb{E}[Y \mid Z^M = 1, i \in \mathcal{C}_j^M] \pi(\mathcal{C}_j^M) + \mathbb{E}[Y \mid Z^M = 1, i \in \mathcal{N}\mathcal{C}_j^M] \pi(\mathcal{N}\mathcal{C}_j^M) \tag{16} \\
&= \mathbb{E}[Y_{j+1} \mid Z^M = 1, i \in \mathcal{C}_j^M] \pi(\mathcal{C}_j^M) + \mathbb{E}[Y_j \mid Z^M = 1, i \in \mathcal{N}\mathcal{C}_j^M] \pi(\mathcal{N}\mathcal{C}_j^M) \tag{17} \\
&= \mathbb{E}[Y_{j+1} \mid i \in \mathcal{C}_j^M] \pi(\mathcal{C}_j^M) + \mathbb{E}[Y_j \mid i \in \mathcal{N}\mathcal{C}_j^M] \pi(\mathcal{N}\mathcal{C}_j^M) \tag{18}
\end{aligned}$$

where Equation 14 follows from the law of total probability, Equation 15 simply substitutes the definitions of $\pi(\mathcal{C}_j^M)$ and $\pi(\mathcal{N}\mathcal{C}_j^M)$, Equation 16 follows from the fact that $\mathcal{S}_j^M = \mathcal{C}_j^M \sqcup \mathcal{N}\mathcal{C}_j^M$, Equation 17 follows from the definition of \mathcal{S}_j^M , and Equation 18 follows from Assumption A3. The same steps generate the equation

$$\mathbb{E}[Y \mid Z^M = 0, i \in \mathcal{S}_j^M] = \mathbb{E}[Y_j \mid i \in \mathcal{C}_j^M] \pi(\mathcal{C}_j^M) + \mathbb{E}[Y_j \mid i \in \mathcal{N}\mathcal{C}_j^M] \pi(\mathcal{N}\mathcal{C}_j^M)$$

and justify the denominator in Equation 13.

Plainly stated, this is a typical LATE estimator where compliance is defined not by voucher use, but by use of the voucher to achieve a specific upward move. Due to random assignment of vouchers, we can assume that non-compliers -regardless of whether they would use an offered voucher or not - are balanced across voucher groups. Note that this identification strategy is predicated on identifying the V_i , as it is required for identifying \mathcal{S}_j^M and applying the Wald estimator to households in this set.

3.3 Estimation Algorithm

The econometrician has no choice over the region of characteristics $\mu(X)$ and U_D for which she will estimate j to $j + 1$ LATEs from the MTO program. The region \mathcal{S}_j^M is determined by the way households select into neighborhoods, and the role of the econometrician ex-post is simply to determine the region over which she can identify transition-specific LATEs using the identification strategy detailed in Sections 3.1 and 3.2. The general estimation algorithm is as follows:

Step 1 Estimate the ordered choice model to obtain estimates of $\hat{\mu}(X_i)$ and \hat{U}_{Di} for all i .

Step 2 Using the control group, find an area $\hat{\mathcal{A}}_j \subset \mathcal{M} \times [0, 1]$ such that households with $(\hat{\mu}(X_i), \hat{U}_{Di}) \in \hat{\mathcal{A}}_j$ select into neighborhood quality $D_i = j$. Using the MTO voucher group, find the subset $\hat{\mathcal{A}}_{j,j+1}^M$ for which households select into neighborhoods of quality $D_i = j$ and $D_i = j + 1$. The identification support set is $\hat{\mathcal{S}}_j^M \equiv \hat{\mathcal{A}}_j \cap \hat{\mathcal{A}}_{j,j+1}^M$.

Step 3 Estimate the j to $j + 1$ transition-specific LATE over $\hat{\mathcal{S}}_j^M$ using the Wald estimator from Equation 13 applied to $\hat{\mathcal{S}}_j^M$.

Step 4 Bootstrap by repeating the following steps T times:

Step 4a Sample with replacement.

Step 4b Repeat Step 1: Estimate the ordered choice model on the new sample.

Step 4c Repeat Step 3: Calculate $\mathbb{E}[\hat{\Delta}_{j,j+1}^{LATE}(Z^M) | (\hat{\mu}(X_i), \hat{U}_{Di}) \in \hat{\mathcal{S}}_j^M]$ on the new sample where the set $\hat{\mathcal{S}}_j^M$ maintains the definition determined in Step 2 for the original sample.

Construct standard errors using the T parameter estimates.

3.4 Comparison with Related Identification Strategies

Appendix D discusses related identification strategies in depth. Here we simply compare our identification strategy to the two most closely related strategies. Closer to our two binary instruments, Angrist and Imbens (1995) show how one can use a binary instrument to identify the Average Causal Response (ACR) parameter, which is a weighted average of transition-specific LATEs. This parameter can be identified simply by applying the Wald estimator with the binary MTO voucher indicator as an instrument. Closer to our transition-specific LATE, Heckman et al. (2006) develop the theory for the case in which one has transition-specific instruments. A complementary approach is to estimate an unordered choice model under restrictions rendering the model estimable (Pinto (2014)).

Vytlacil (2006) shows that Angrist and Imbens (1995)'s assumptions are equivalent to an ordered choice model with random cutoffs. We must adopt the stronger assumption of constant cutoffs, but in exchange we are able to identify transition-specific parameters without any transition-specific instruments.

4 Moving to Opportunity (MTO)

4.1 Program Description

Moving To Opportunity (MTO) was inspired by the promising results of the Gautreaux housing mobility 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)).¹⁰

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 continued receiving project-based assistance.¹¹

¹⁰It 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)).

¹¹Section 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.

4.2 Data

The first source of data we use in our analysis is the MTO Interim Evaluation. The MTO Interim Evaluation 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.

4.2.1 Variables

We create a variable measuring neighborhood quality using a linear combination of several neighborhood characteristics. Neighborhood characteristics measured by NHGIS variables are first transformed into percentiles of the national distribution from the 2000 census. Principal components analysis is then used to determine which single vector accounts for the most variation in 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 Employed-to-Population Ratio (EPR), and the female unemployment rate.¹²

The resulting univariate index explains 69 percent of the variance of these neighborhood characteristics, and Table 1 reports that no additional eigenvector would explain more than 11 percent of the variance of these variables. Table 1 also displays the coefficients relating each of these variables to the index vector, the magnitudes of which are similar to the magnitude of the coefficient for poverty for most variables. Finally, while poverty is negatively correlated with quality as expected, Figure 2 shows the existence near the 10 percent poverty cutoff of eligibility for using an MTO voucher, which is approximately the median of the national distribution, of neighborhoods of both very high and very low quality. This issue will be revisited when examining sorting patterns in Section 5.2.

¹²Given the importance of neighborhood violence (Anderson (1999), Aliprantis (2014)), especially as documented in the context of MTO (Kling et al. (2005)), we would like to include measures of neighborhood violence and the rule of law. We do not include such measures in our index of quality because to the authors' knowledge variables comparable to those used in this analysis are not available.

Table 1: 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

Note: 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.

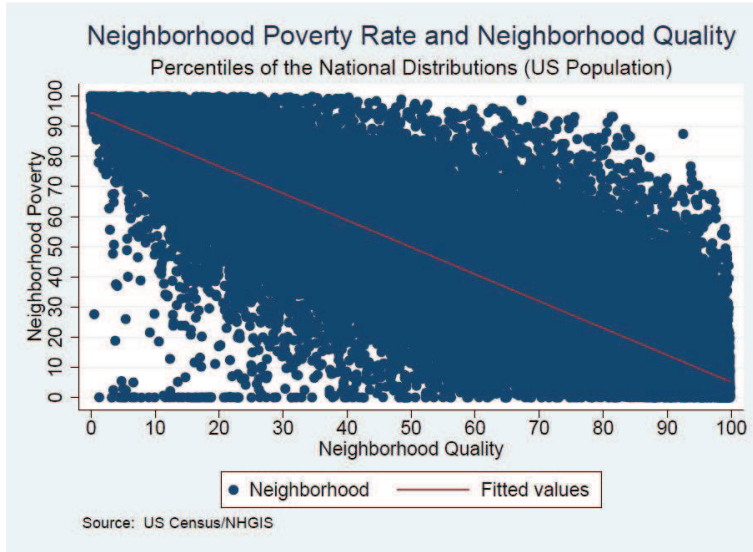


Figure 2: Neighborhood Poverty and Quality

Note: 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.

Quality of the baseline and 2002 neighborhood of residence is measured using 2000 Census data. Our static model has some implicit assumptions about dynamics, such as how long one must stay in a neighborhood for the move to contribute to one’s treatment, and we believe these assumptions are best addressed using 2000 Census data.

In addition to baseline neighborhood quality, other baseline characteristics of the MTO households used in this model are whether the respondent had family living in their neighborhood of residence, whether a member of the household was a victim of a crime in the previous 6 months, and whether there were teenage children in the household. Site of residence is the only other observed characteristics included in X_i ; when models were estimated with additional variables like number

of children or residence in an early HOPE VI building, the coefficients on these other observables were all statistically insignificant.

Outcome variables for adults from the MTO Interim Evaluation include the labor market status of the adult at the time of the interim survey (ie, Two binary variables, one indicating labor force participation, the other indicating whether the adult was employed.), the self-reported total household income (all sources summed), the individual earnings in 2001 of the sample adult, receipt of Temporary Assistance for Needy Families (TANF) benefits, and the respondent’s body mass index (BMI). Weights are used in constructing all estimates.¹³

4.2.2 Sample and Descriptive Statistics

The focus of this analysis is adults in the MTO Interim Evaluation sample. We choose not to include children in our analysis for two reasons. First, careful analysis of program effects from MTO on children’s outcomes has already been conducted (Sanbonmatsu et al. (2006), Kling et al. (2007)). Second, as shown in Figure 3 and investigated in depth in DeLuca and Rosenblatt (2010), MTO did not induce large changes in school quality.¹⁴ Both intuition and previous findings in the literature suggest schools may be the most relevant “neighborhood” for children (Oreopoulos (2003)).

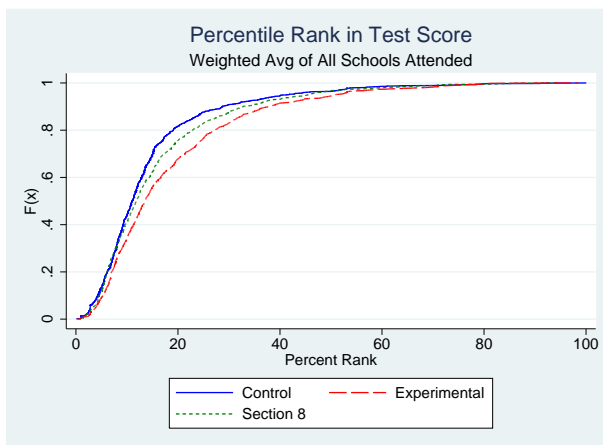


Figure 3: School Ranking on State Tests, Weighted Average Percentile over all Schools Attended
 Note: This Figure shows Cumulative Distribution Functions (CDFs) of the percentile rank of the schools attended by MTO children on state-level standardized test scores. 0 is the lowest ranked school in the state, and 100 is the highest ranked school in the state.

The ensuing analysis is focused on a sample that is restricted in two ways. The first restriction ensures that we are focusing on a relatively homogeneous population. To satisfy this restriction

¹³Weights 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 sub-sampling procedures used during the interim evaluation data collection.

¹⁴The advent of school choice may be important for these results: Thirty percent of MTO control group children in Chicago and Los Angeles were attending magnet schools (Sanbonmatsu et al. (2006), p 684).

we drop all households living at baseline in a neighborhood above the tenth percentile of the national distribution of quality. Looking at Figure 4b, we can see that the median baseline neighborhood quality for MTO participants was below the first percentile of the national distribution. For Chicago, Los Angeles, and New York City, nearly all participants lived at baseline in neighborhoods below the 10th percentile of the national distribution. In Baltimore and Boston, however, at baseline a non-trivial share of program participants lived in higher quality neighborhoods, driven mainly by the male EPR and the share of adults holding a BA in their neighborhoods. These individuals represent a little under 15 percent of the interim evaluation sample, and are dropped from our estimation sample.

The second sample restriction facilitates the estimation of the ordered choice model. To satisfy this restriction we top-code neighborhood quality at the the median of the national distribution of quality in 2000. Figure 4d shows the final results of these restrictions, with Figures 4b and 4c showing the original sample and the one resulting from the first restriction alone, respectively.

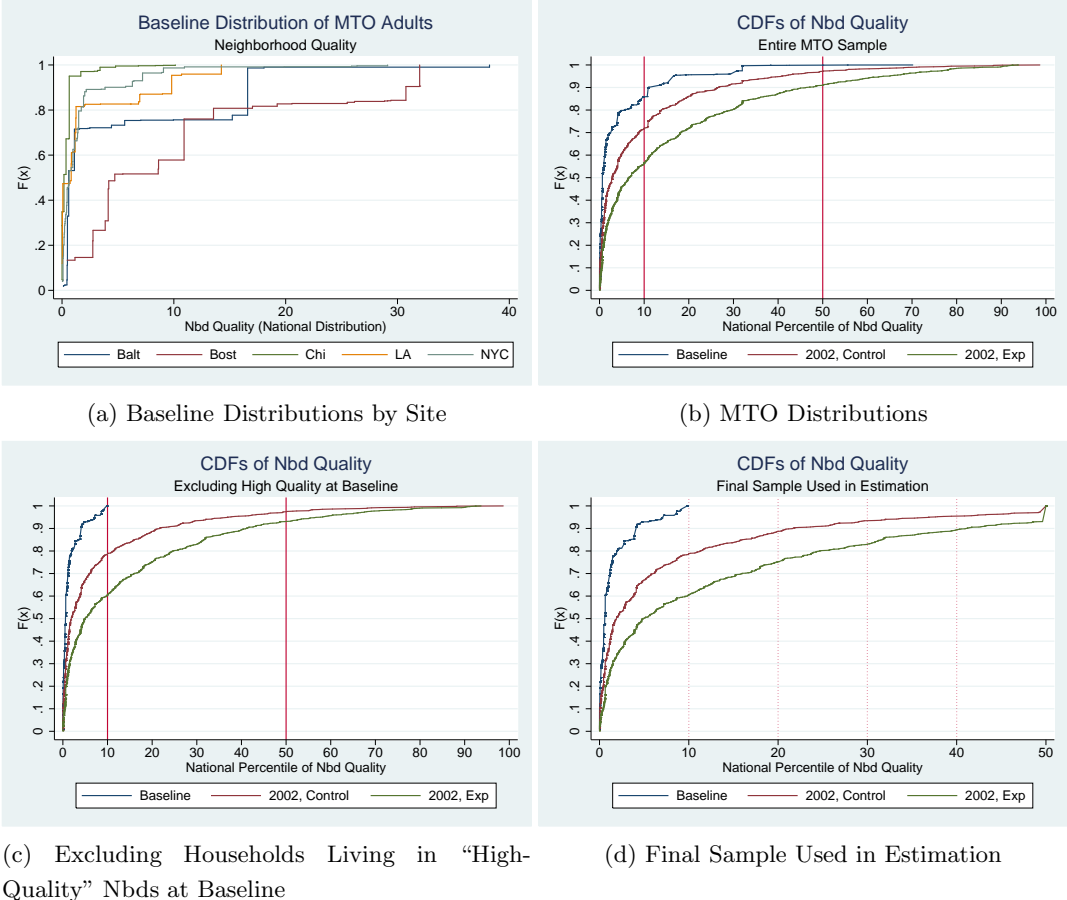


Figure 4: Neighborhood Quality in MTO

Note: This Figure shows the neighborhood quality distribution of MTO adults in various subsamples.

The final estimation sample used in our analysis has a little under 3,100 adults (a little over 85 percent of the interim sample and a little under 75 percent of the original adult sample). Our

sample represents “the other one percent:” At baseline, 67 percent of the estimation sample lived in a neighborhood whose quality was below the 1st percentile of the national distribution of neighborhood quality. There was enough mobility in the control group so that by 2002 this had decreased to 39 percent living in first percentile neighborhoods. On the other hand, though, the mobility of the control group was not extraordinary. By 2002, about 80 percent of the sample in the control group lived in a neighborhood whose quality was less than the 10th percentile of the national distribution of neighborhood quality.

Although we include Boston in the analysis, it is important to note that it is clearly an outlier relative to the other MTO sites. In Figure 4a we see that unlike all of the other sites, the baseline neighborhood quality in Boston was not confined to the first percentiles of the national distribution of neighborhood quality. Table 2 quantifies these differences precisely at the time of the interim evaluation. We can see that the control group in Boston looks more like the experimental group in every other site. Boston comprised approximately 15 percent of the weighted observations in the estimation sample.

Figure 4b and Table 2 reproduce the broad result from Aliprantis (2015a) that MTO induced small shares of adult participants into high quality neighborhoods.

Table 2: Interim Neighborhood Quality by Site

Site	Mean			Median		
	Control	Section 8	Experimental	Control	Section 8	Experimental
Baltimore	8.8	12.8	15.1	2.2	4.4	4.9
Boston	15.1	20.3	25.9	10.9	13.5	17.0
Chicago	8.3	11.2	12.5	2.9	3.6	3.9
Los Angeles	6.0	7.6	13.2	1.2	3.9	6.3
New York City	5.7	8.2	13.7	1.0	1.7	2.6

5 Empirical Specification and Estimation Results

5.1 Ordered Choice Model Specification

When estimating the parameters of the ordered choice model, we define the discrete treatment levels by the intervals $\{\underline{q}_k^Q, \bar{q}_k^Q\}$ dividing the estimation sample into its deciles at the time of the

interim survey

$$Q_i = \begin{cases} 1 & \text{if } q_i \in [\underline{q}_1^Q, \overline{q}_1^Q) = [0, 0.2); \\ 2 & \text{if } q_i \in [\underline{q}_2^Q, \overline{q}_2^Q) = [0.2, 0.6); \\ 3 & \text{if } q_i \in [\underline{q}_3^Q, \overline{q}_3^Q) = [0.6, 1.1); \\ 4 & \text{if } q_i \in [\underline{q}_4^Q, \overline{q}_4^Q) = [1.1, 1.9); \\ 5 & \text{if } q_i \in [\underline{q}_5^Q, \overline{q}_5^Q) = [1.9, 3.6); \\ 6 & \text{if } q_i \in [\underline{q}_6^Q, \overline{q}_6^Q) = [3.6, 6.1); \\ 7 & \text{if } q_i \in [\underline{q}_7^Q, \overline{q}_7^Q) = [6.1, 11); \\ 8 & \text{if } q_i \in [\underline{q}_8^Q, \overline{q}_8^Q) = [11, 19); \\ 9 & \text{if } q_i \in [\underline{q}_9^Q, \overline{q}_9^Q) = [19, 33); \\ 10 & \text{if } q_i \in [\underline{q}_{10}^Q, \overline{q}_{10}^Q) = [33, 50). \end{cases}$$

The marginal benefit of choosing to move from treatment level k to $k + 1$ in the ordered choice model from Section 3.1 is specified to be

$$MB_{ik} = \mu(X_i) + \gamma_k^S Z_i^S \tau_i^S + \gamma_k^M Z_i^M \tau_i^M - C_k - V_i.$$

We specify the components of MB_{ik} to be

$$\begin{aligned} \mu(X_i) &= \beta_1 X_1 + \dots + \beta_8 X_8 \\ \gamma_k^S &= \Gamma_0^S + \Gamma_k^S \\ \gamma_k^M &= \Gamma_0^M + \Gamma_k^M \\ C_k &= \delta_0 + \delta_k, \end{aligned}$$

with households' decisions to move with a voucher determined by the similarly-specified latent index models:

$$\begin{aligned} \tau_i^S &= \mathbf{1}\{ \mu^S(X_i) - V_i^S \geq 0 \} = \mathbf{1}\{ \beta_1^S X_1 + \dots + \beta_8^S X_8 - V_i^S \geq 0 \}, \\ \tau_i^M &= \mathbf{1}\{ \mu^M(X_i) - V_i^M \geq 0 \} = \mathbf{1}\{ \beta_1^M X_1 + \dots + \beta_8^M X_8 - V_i^M \geq 0 \}. \end{aligned}$$

Recall that V_i represents the unobserved cost for household i of moving up one level in the absence of a voucher program, and V_i^S and V_i^M are unobserved variables influencing the decision of household i to take up a Section 8 voucher and an MTO voucher when these are offered. We allow for these variables to be correlated in an arbitrary way, possibly exhibiting patterns of correlation anywhere between being exactly identical variables to being independently distributed variables to being negatively correlated variables. However, for the sake of identification we do adopt the distributional assumption:

A6:

$$\mathbf{V}_i \equiv (V_i, V_i^S, V_i^M) \sim \mathcal{N} \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho^S & \rho^M \\ \rho^S & 1 & \rho^{SM} \\ \rho^M & \rho^{SM} & 1 \end{bmatrix} \right).$$

We stress that the role of Assumption A6 in aiding identification is entirely through the choice model. We also stress that while the normal distribution of V_i may seem like the imposition of a functional form, it is a normalization when the specification of $C(q)$ is flexible (See the discussion in Footnote 8.). The full likelihood is specified in Appendix B, including a discussion of why ρ^S and ρ^M are sufficient to identify the choice model parameters even without identifying ρ^{SM} .

Recall that the first four variables in X_i are baseline neighborhood quality, whether the respondent had family living in the neighborhood of residence, whether a member of the household was a victim of a crime in the previous 6 months, and whether there were teenage children in the household at baseline. The final four variables in X_i are site indicators. Γ_0^S and Γ_0^M are themselves site-specific fixed effects that capture differences in factors like local labor and housing markets across sites. We do not attempt to explicitly model housing market prices since these are largely offset by the nature of the payment structure of the rental vouchers and project-based programs. Individuals pay 30 percent of their income towards rent in project-based units, and pay this same rent when using vouchers as long as the price of rent is not above the FMR. Thus, the probability of leasing up with a Section 8 voucher is more salient than price; this could change if FMRs were determined for finer geographies than metro areas (Collinson and Ganong (2013)). Further discussion of the interpretation of model parameters is provided in Appendix E.

Like Galiani et al. (2015), we interpret Z^S and Z^M as the random assignment of a potential reduction in the cost to accessing a higher quality neighborhood relative to staying in the baseline neighborhood. It is important to note that secular trends outside the control of the program might swamp this cost reduction. One can imagine changing costs to accessing higher quality neighborhoods due to changes in the local labor or housing markets, changes in school quality due to the provision of magnet schools, or simply an improvement in the quality of the baseline neighborhood.

We estimate the parameters of this ordered choice model via Maximum Likelihood using the log-likelihood function in Appendix B. We estimate \widehat{V}_i as in Equation 12 after linearly interpolating between the \widehat{C}_k and $\widehat{\gamma}_k$ as in Equations 6 and 10.

5.1.1 Potential Outcomes Specification

When estimating treatment effect parameters we define the discrete treatment levels in the intervals $\{q_j^D, \bar{q}_j^D\}$ in terms of deciles of the national distribution:

$$D_i = \begin{cases} 1 & \text{if } q_i \in [q_1^D, \bar{q}_1^D) = [0, 10); \\ 2 & \text{if } q_i \in [q_2^D, \bar{q}_2^D) = [10, 20); \\ 3 & \text{if } q_i \in [q_3^D, \bar{q}_3^D) = [20, 30); \\ 4 & \text{if } q_i \in [q_4^D, \bar{q}_4^D) = [30, 40); \\ 5 & \text{if } q_i \in [q_5^D, \bar{q}_5^D) = [40, 50). \end{cases}$$

We choose deciles to discretize neighborhood quality when investigating treatment effects not because we believe treatment should have an effect when crossing the particular thresholds of neighborhood quality used in this definition, but because we believe it offers the best balance between theoretical ideal and practical necessity. The model assumes that moves within a given level of treatment will not have effects on outcomes. Even if they do, it is enough to assume that individuals do not select *within* treatment levels based on rich information regarding neighborhood quality.¹⁵ If these assumptions do not hold within entire deciles of quality, the effects from such moves will likely enter the estimation results through the U_j . Theoretically, one way to handle this issue would be to increase the number of bins until moves within a given level do not have effects on outcomes. Another way to handle this problem would be to reformulate the model to accommodate a continuous treatment (Florens et al. (2008)).

Due to the limited mobility induced by MTO, we believe deciles of quality offer the smallest window on which it is feasible to estimate neighborhood effects using the MTO interim survey data. As the next Section shows, even this discretization leaves us with undesirably small sample sizes of compliers, resulting in little power to detect even very large effects. As a result, the only LATEs we attempt to estimate are of moves between $j = 1$ and $j = 2$.

5.2 Estimation Results: Ordered Choice Model

Figure 5 shows the model fit as characterized by the estimated distributions of observed and unobserved variables. It is important to note that of the distributions shown in Figure 5, only Figure 5c is weighted.

¹⁵This can be seen as a stronger version of the central identifying assumption in Bayer et al. (2011).

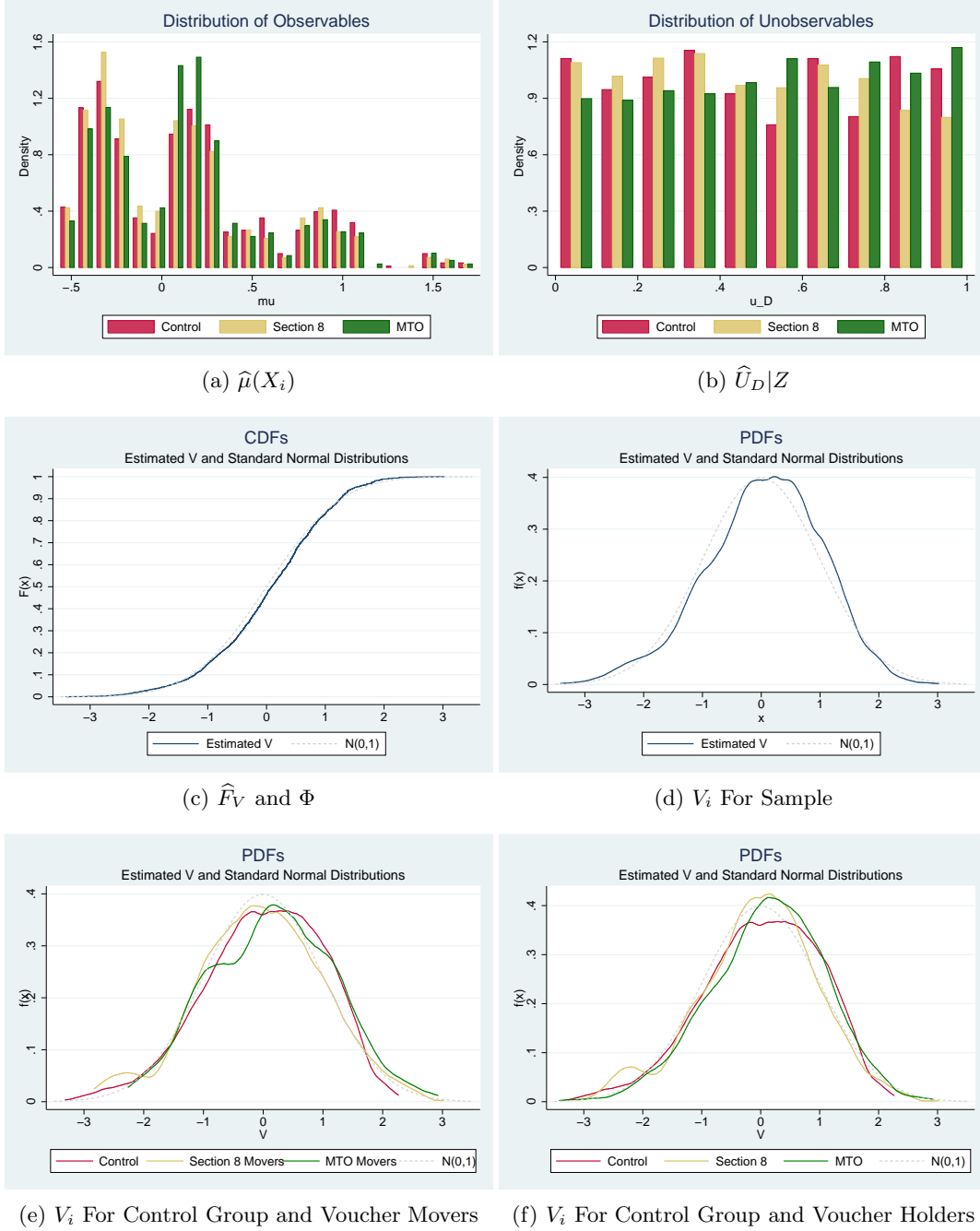


Figure 5: Model Fit: The Distributions of Observables and Unobservables

Note: This Figure helps to illustrate the fit of the neighborhood choice model. (c) and (d) are useful in assessing whether the estimated \hat{V}_i appear to follow a normal distribution; (b), (e), and (f) are useful in assessing whether the estimated \hat{V}_i are systematically different across voucher groups or groups of movers; and (a) is useful in assessing whether the observed component of the estimated choice model is systematically different across voucher groups. Only (c) is weighted.

The estimated cost function $\hat{C}(q)$ is shown in Figure 6 together with the cutoffs for treatment in the ordered choice model ($\{\underline{q}_k^Q, \bar{q}_k^Q\}$) and for effects on outcomes ($\{\underline{q}_j^D, \bar{q}_j^D\}$). The cost function is estimated to take the expected shape.

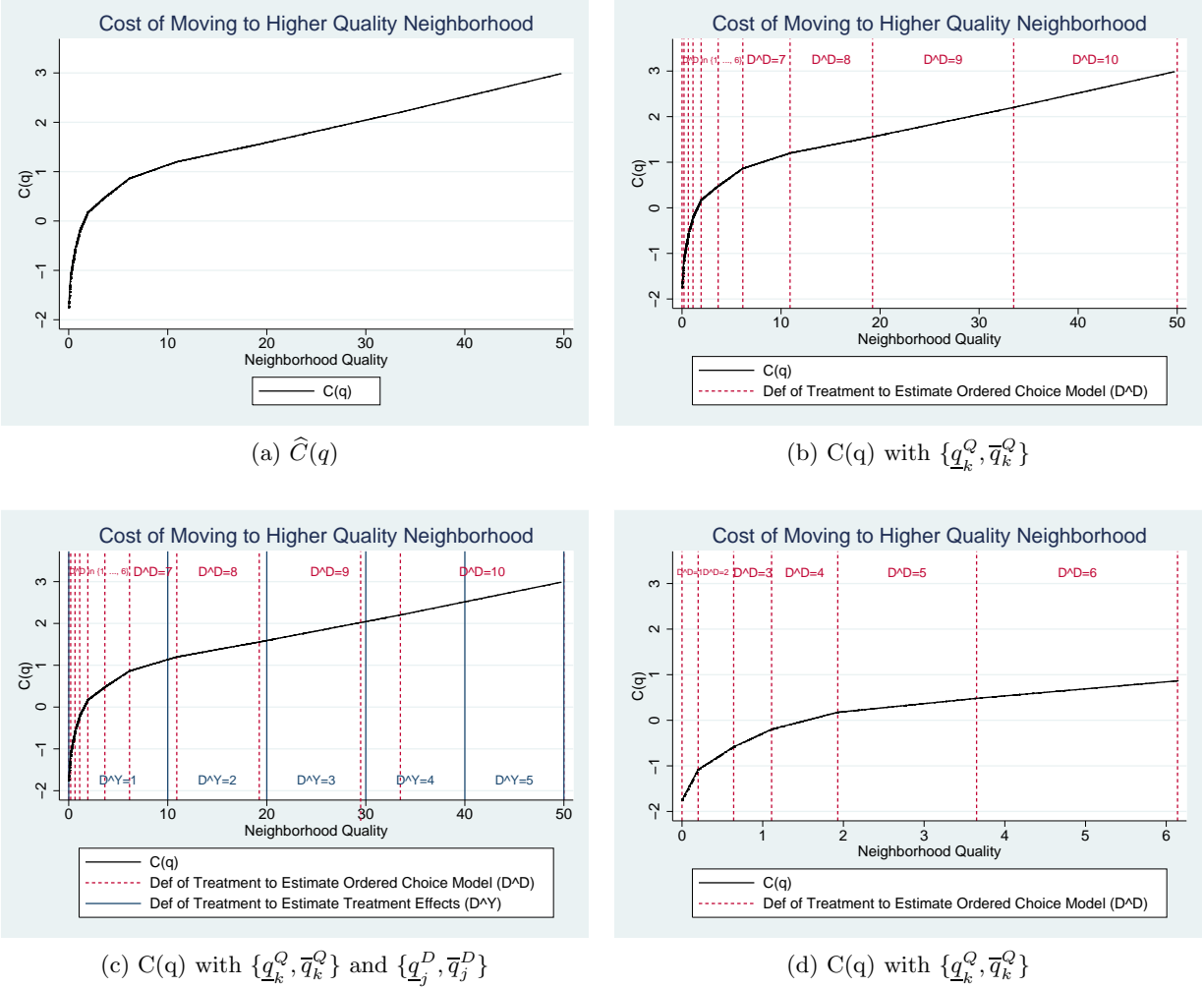
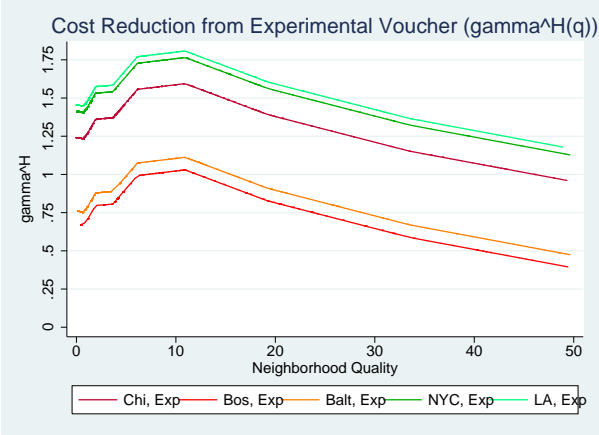


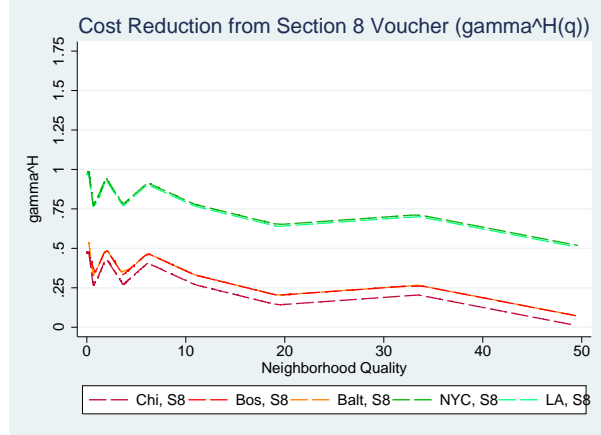
Figure 6: Cost Function

Note: (a) shows the estimated cost function $\hat{C}(q)$. The dashed red vertical lines in (b) show the discretization of quality used to define Q_i , and the cost function $C(q)$ is constructed by linearly interpolating between the cutpoints estimated in an ordered probit model. The solid blue vertical lines in (c) show the discretization of quality used to define D_i as used in the model of potential outcomes. (d) zooms in to show the detail of (b) since most of the estimation sample was in low-quality neighborhoods at the time of the interim evaluation.

Several stylized facts emerge from the estimated cost reductions $\hat{\gamma}^S(q)$ and $\hat{\gamma}^M(q)$ displayed in Figure 7, allowing us to characterize the Section 8 and experimental MTO voucher programs. First, the MTO voucher was much more effective than the standard Section 8 voucher in getting complier households to access higher quality neighborhoods. Second, the effectiveness of both types of vouchers varied considerably by site. Vouchers represented the largest cost-reductions in Los Angeles (LA) and New York City (NYC), and represented the smallest cost-reductions in Baltimore and Boston. Chicago displays the largest gap in cost reduction between programs.



(a) $\gamma^M(q)$ - Experimental Voucher Cost Reduction



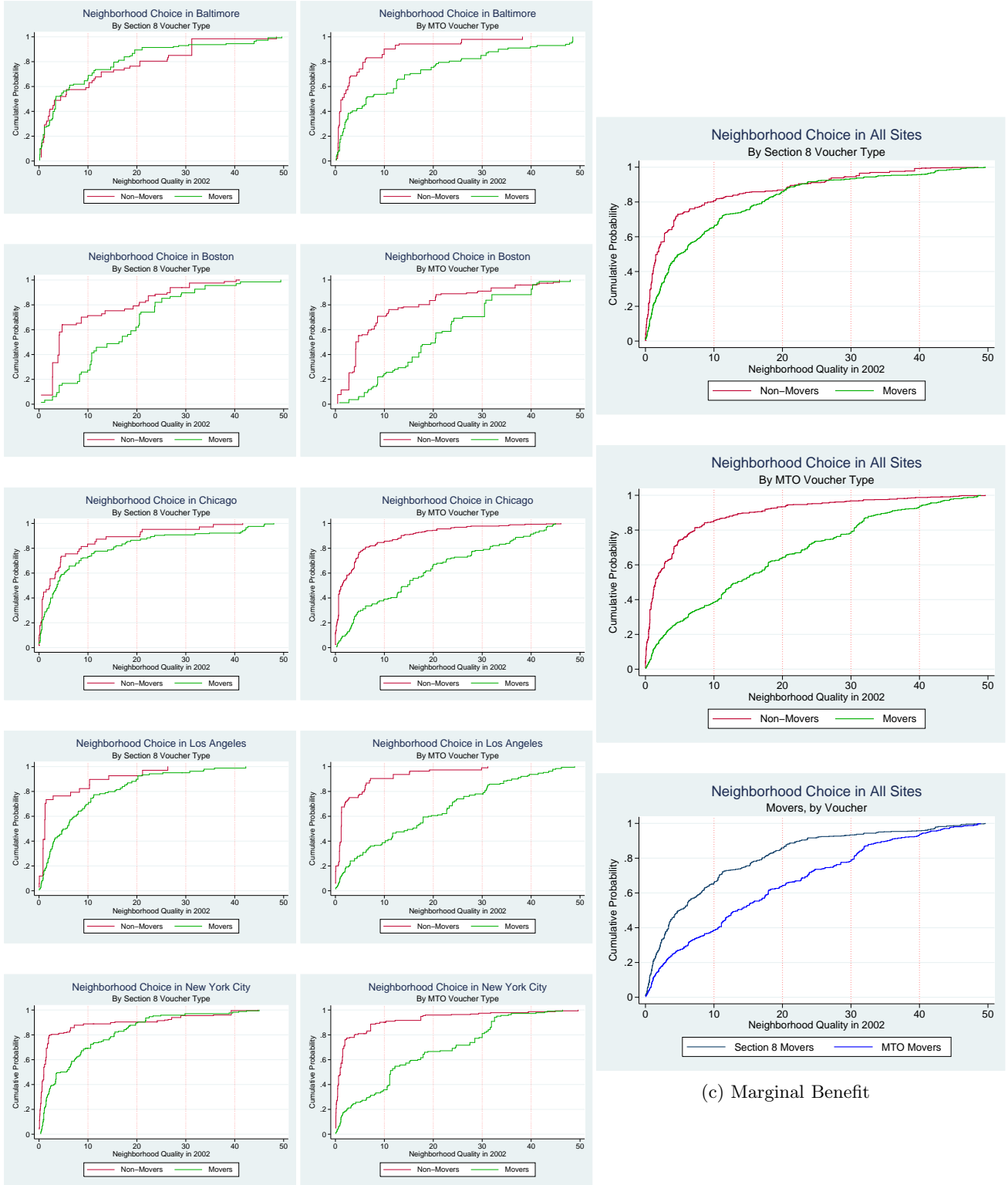
(b) $\gamma^S(q)$ - Section 8 Voucher Cost Reduction

Figure 7: Cost Reduction

Note: (a) shows the estimated cost reduction when moving with an experimental MTO voucher, and (b) shows the estimated cost reduction when moving with a standard Section 8 voucher. A higher cost reduction implies the voucher is more effective at the site in helping residents to move to a higher quality neighborhood.

Figure 8 compares the neighborhood quality of mover and non-mover households after receiving vouchers. For Section 8 voucher holders in Chicago and Baltimore, we see that the neighborhood quality of movers is almost indistinguishable from that of non-movers, while for NYC and LA improvements in neighborhood quality are small.¹⁶ Boston is the only site in which movers with a Section 8 voucher resided in significantly higher quality neighborhoods than non-movers. Across all sites, MTO compliers are much more likely than Section 8 compliers to access neighborhoods with a quality index above the first decile.

¹⁶See Ludwig et al. (2005) for a study of MTO's program effects in Baltimore.



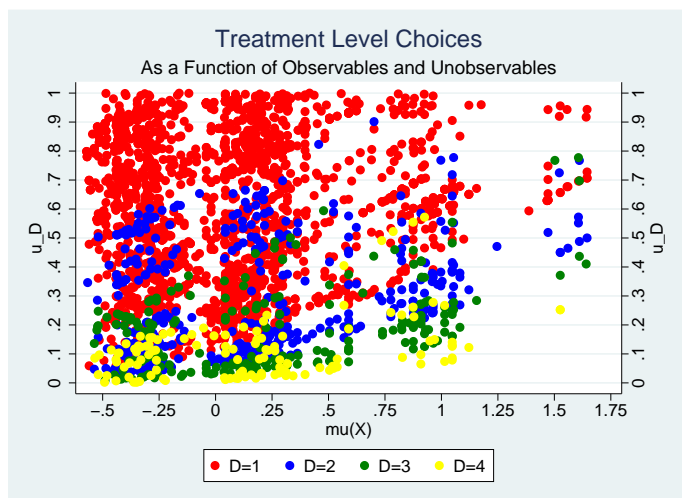
(a) Section 8 Voucher Recipients (b) MTO Voucher Recipients

(c) Marginal Benefit

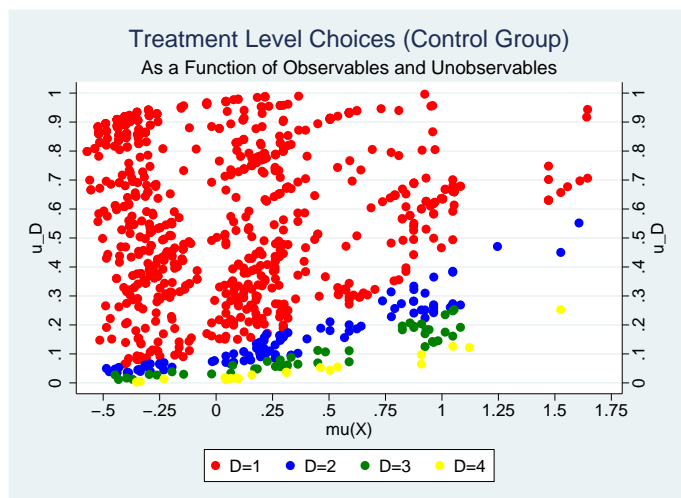
Figure 8: Selection into Neighborhood Quality by Voucher and Type

Note: This Figure shows raw data consistent with the estimated neighborhood choice model, showing in (c) that the gap in neighborhood quality between households who moved with an MTO voucher was much larger than the gap in quality between households who moved with a standard Section 8 voucher. Figures (a) and (b) show that this pattern is stable across sites.

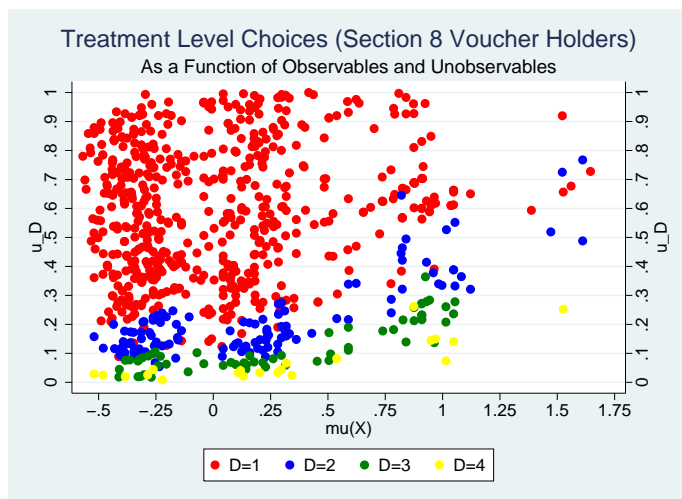
The differences between the Section 8 and MTO voucher programs are evident when examining the actual mobility of program participants as shown in Figure 9. To begin, Figure 9a shows all program participants, color-coded by whether they lived in a neighborhood at the interim evaluation ranked in the first, second, third, or fourth decile of the national distribution of neighborhood quality. On the x -axis is the $\hat{\mu}(X_i)$ of each household, and on the y -axis is the percentile of the household's unobserved determinant of selection in the absence of a program, $\hat{U}_{Di} \equiv F_V(\hat{V}_i) = \Phi(\hat{V}_i)$.



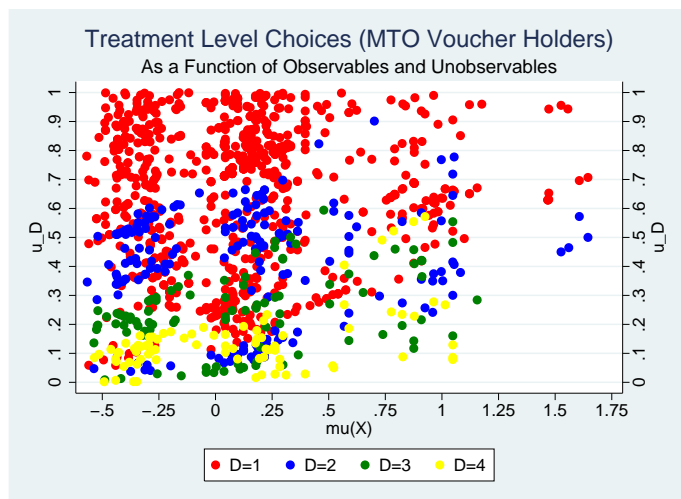
(a) All Groups



(b) Control Group



(c) Section 8 Voucher Holders



(d) MTO Voucher Holders

Figure 9: Selection into Treatment

Note: This Figure shows how MTO adults sorted into the discrete quality levels $D_i \in \{1, 2, 3, 4\}$ as a function of their estimated choice model observed characteristics ($\hat{\mu}(X_i)$ - on the x -axis) and unobserved characteristics (\hat{U}_{Di} - on the y -axis). Since these characteristics determine how households would select into D_i in the absence of any voucher, and vouchers were randomly assigned to households, (b)-(d) characterize counterfactual choices under ideal interventions to voucher status.

Since vouchers were randomly assigned in MTO, Figures 9b-9d illustrate counterfactual distri-

butions of neighborhood quality under external manipulations to voucher type. For each household, given observed variables summarized by $\mu(X_i)$ and unobserved variables u_{Di} , these figures show the neighborhood quality households would select into under each setting of the vouchers.

Proceeding to Figure 9b, we can see that almost all of the control group remained in low-quality neighborhoods, most remaining in the first decile of neighborhood quality. Only households with very high observed factors $\hat{\mu}(X_i)$ and very low unobserved cost factors \hat{U}_D managed to move to higher quality neighborhoods, even when defined as moving only to the second, third, or fourth deciles of the national distribution. In the bottom left corner of Figure 9c some red dots have turned to blue, indicating that for low \hat{U}_D some households would be induced by the Section 8 voucher to move to a higher quality neighborhood. However, the similarity of Figures 9b and 9c is quite remarkable, suggesting that Section 8 vouchers were not very effective in getting households to move to higher quality neighborhoods.

Figure 9d shows that the MTO voucher was far and away more effective in getting complier households to move to higher quality neighborhoods than the ordinary Section 8 voucher. Unfortunately, we see that most of this mobility is from the first to the second decile of the national distribution of quality. Although it was still relatively rare, the MTO voucher did manage to induce some households to move into the third and fourth deciles of neighborhood quality.

The differences between the effects of Section 8 and MTO vouchers on selection into neighborhood quality are even more interesting when one considers that although 59 percent of our sample moved with the voucher when offered a Section 8 voucher, only 43 percent of MTO voucher recipients moved with their voucher. This suggests that simply asking whether recipients take-up a voucher and move when it is offered need not be the best way to judge the effectiveness of housing mobility vouchers. These results also reiterate that selection into neighborhoods of various quality levels is an inherently interesting phenomenon to study (Clampet-Lundquist and Massey (2008), Sampson (2008), Sampson (2012)).

Table 3 reports the estimated coefficients in $\hat{\mu}(X)$ and the site fixed effects in $\hat{\delta}_0$. There are large differences between sites: In the absence of any voucher program, program participants in Boston tended to live in much higher quality neighborhoods than their counterparts in the other MTO sites at the time of the interim evaluation. Baltimore and Chicago tended to be similar, while LA and NYC were the worst sites by far.

At the household level, we see that having teens in the household reduces the likelihood of moving with and without a voucher. It is possible that room occupancy restrictions according to age and gender of children may have made it harder for families with older children to find housing. Having no family in the baseline neighborhood makes households more likely to move. Living in a higher quality neighborhood at baseline increases the likelihood of moving without a voucher, but decreases moving prospects post MTO voucher assignment and counseling. The fact that the correlation between V and V^M is negative suggests that MTO was able to increase the likelihood of moving for households that had intrinsically larger unobserved costs to move.

Table 3: Ordered Choice Model Parameter Estimates

X_k and \mathbf{V}	$\hat{\beta}_k$	$\hat{\beta}_k^S$	$\hat{\beta}_k^M$
Baseline Characteristics			
Teens in HH	-0.08 (0.05)	-0.48 (0.10)	-0.39 (0.09)
Family in Nbd	-0.14 (0.05)	-0.16 (0.12)	0.00 (0.06)
HH Member Victim	0.03 (0.05)	0.10 (0.10)	0.10 (0.10)
Baseline Nbd Quality	0.13 (0.01)	-0.02 (0.03)	-0.10 (0.02)
Site Fixed Effects/Constant			
Baltimore	0 -	0.03 (0.13)	-0.25 (0.13)
Boston	0.31 (0.10)	-0.49 (0.21)	0.02 (0.02)
Chicago	-0.04 (0.09)	0.04 (0.13)	-0.50 (0.13)
Los Angeles	-0.52 (0.10)	0.39 (0.17)	0.47 (0.13)
New York City	-0.58 (0.09)	0.60 (0.15)	-0.13 (0.15)
Unobserved Factors			
ρ^S and ρ^M	- -	0.07 (0.10)	-0.17 (0.11)

Note: The first column reports parameters of $\mu(X)$ corresponding to the choice model in the absence of any program, identified by the control group. The second column reports parameters of $\mu^S(X)$, which determines how likely households are to take-up a Section 8 voucher. The third column analogously reports parameters of $\mu^M(X)$. ρ^S and ρ^M are the correlations between the unobserved component determining neighborhood choice without a program and the unobserved component determining the likelihood of take-up with a voucher of either program type. The full likelihood is specified in Appendix B.

5.3 Estimation Results: LATEs of Neighborhood Quality

5.3.1 What Effects are Identified?

Recalling the counterfactual distributions displayed in Figure 9, there is a range of values of $(\mu(X_i), U_{Di})$ in $\mathcal{M} \times [0, 1]$ for which households would be induced by receiving an MTO voucher to move from a $D = 1$ quality neighborhood to quality $D = 2$.¹⁷ There is another range for which households would not move from $D = 1$ (those with high U_{Di}), and there are also ranges for which households would be induced to make other moves, such as from $D = 2$ to $D = 3$.

¹⁷We refer interchangeably to parameters and sets defined in terms of V_i and U_{Di} , where $U_{Di} = F_V(V_i)$.

Due to the observed patterns of neighborhood selection displayed in Figure 9, we focus on identifying effects of moving from the first to the second decile of neighborhood quality.¹⁸ Table 4 characterizes some of the changes in neighborhood characteristics that would typically accompany a move from $D = 1$ to $D = 2$. On average, the poverty rate would decline from 33 to 22 percent, BA attainment would go from 7 to 11 percent, the share of single-headed households would drop from 52 to 38 percent, and the female unemployment rate would drop from 16 to 10 percent. While these changes in neighborhood characteristics are non-trivial, it is worth pointing out that they are still far worse than the unconditional median neighborhood in the US in 2000, and changes of these magnitudes would have to occur several times to achieve the characteristics of the highest quality neighborhoods. As discussed in Aliprantis (2015a) and elsewhere, these are moves from the most extreme areas of the left tail of the distribution of quality to neighborhoods that are still within the left tail of quality.

Table 4: Average Neighborhood Characteristics in 2000 Conditional on Neighborhood Quality

Nbd Characteristic	Mean $ D = 1$	Mean $ D = 2$	Unconditional Median	Mean $ D = 10$
Poverty Rate (%)	33	22	9	3
HS Diploma (%)	55	65	83	95
BA (%)	7	11	19	52
Single-Headed HHs (%)	52	38	24	11
Female Unemployment Rate (%)	16	10	5	2
Male EPR ($\times 100$)	55	65	79	89

This characterization of the neighborhood mobility induced by the MTO voucher might be surprising if we were to define neighborhood quality in terms of poverty alone. Recall Figure 2 and the discussion in Section 4.2: There exist low quality neighborhoods that are also low-poverty. Table 5 quantifies the prevalence of these neighborhoods in MTO states in 2000. The Table shows that while moving with an MTO voucher essentially ruled out neighborhoods in the lowest decile of quality, MTO voucher holders had many options for using their vouchers in low-quality neighborhoods that met the MTO voucher poverty restriction. Additionally, households moving with MTO experimental vouchers could end up in low-quality neighborhoods because the voucher only had to be used for one year before residents could use it move again without any restrictions, or because rent increases made it impossible for them to stay in their new locations (de Souza Briggs et al. (2010)).

¹⁸We have also estimated Average Causal Responses (ACRs) from Angrist and Imbens (1995) for subsets in which many possible moves are induced. The results are broadly consistent with our LATE estimates.

Table 5: Low-Poverty ($\leq 10\%$), Low-Quality ($D \leq 3$) Neighborhoods in MTO States in 2000

Nbd Quality	Number of Residents
$D = 1$	6,362
$D = 2$	93,385
$D = 3$	751,738

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

5.3.2 For Whom are Effects Identified?

We proceed graphically, using Figure 10 to empirically implement the procedure from Section 3.3 to determine the support of $(\mu(X_i), U_{Di})$ for which LATEs of moving from $D = 1$ to $D = 2$ are identified. We can define the area \mathcal{A}_1 for which households would select into neighborhood quality $D = 1$ without any voucher (Figure 10h). We can also define the subset $\mathcal{A}_{1,2}^M$ for which households would select into neighborhood quality $D = 2$ with an MTO voucher (Figure 10i). Defining $\mathcal{S}_{1,2}^M \equiv \mathcal{A}_1 \cap \mathcal{A}_{1,2}^M$, the identification support set is

$$\mathcal{S}_{1,2}^M \equiv \left\{ (\mu(X_i), U_{Di}) \mid \mu(X_i) \in [-0.6, 0.4], \quad U_{Di} \in [0.43+0.30\mu(X_i), 0.68+0.15\mu(X_i)] \right\},$$

which can be seen in Figures 10d-10f.

5.3.3 LATEs of Neighborhood Quality Estimation Results

Estimates of LATEs of neighborhood quality are reported for the subpopulation of compliers in $\mathcal{S}_{1,2}^M$ in Table 6. All of these effects conform with the theory that living in higher quality neighborhoods improves adult labor market and health outcomes while decreasing receipt of welfare benefits. All of these point estimates are large, and even moving from the first to second levels of neighborhood quality alone is estimated to have statistically significant effects at the 10 percent level or smaller on adult labor market outcomes like labor force participation, employment, and household income. A statistically significant effect is also found on welfare receipt, which is estimated to be entirely eliminated. Although the estimated effect on Body Mass Index (BMI) is large in magnitude, it is not statistically significant.

While some of the LATEs in Table 6 are imprecisely estimated, it is difficult to interpret any of the estimated effects as evidence against the theory that living in higher quality neighborhoods improves adult labor market and health outcomes.

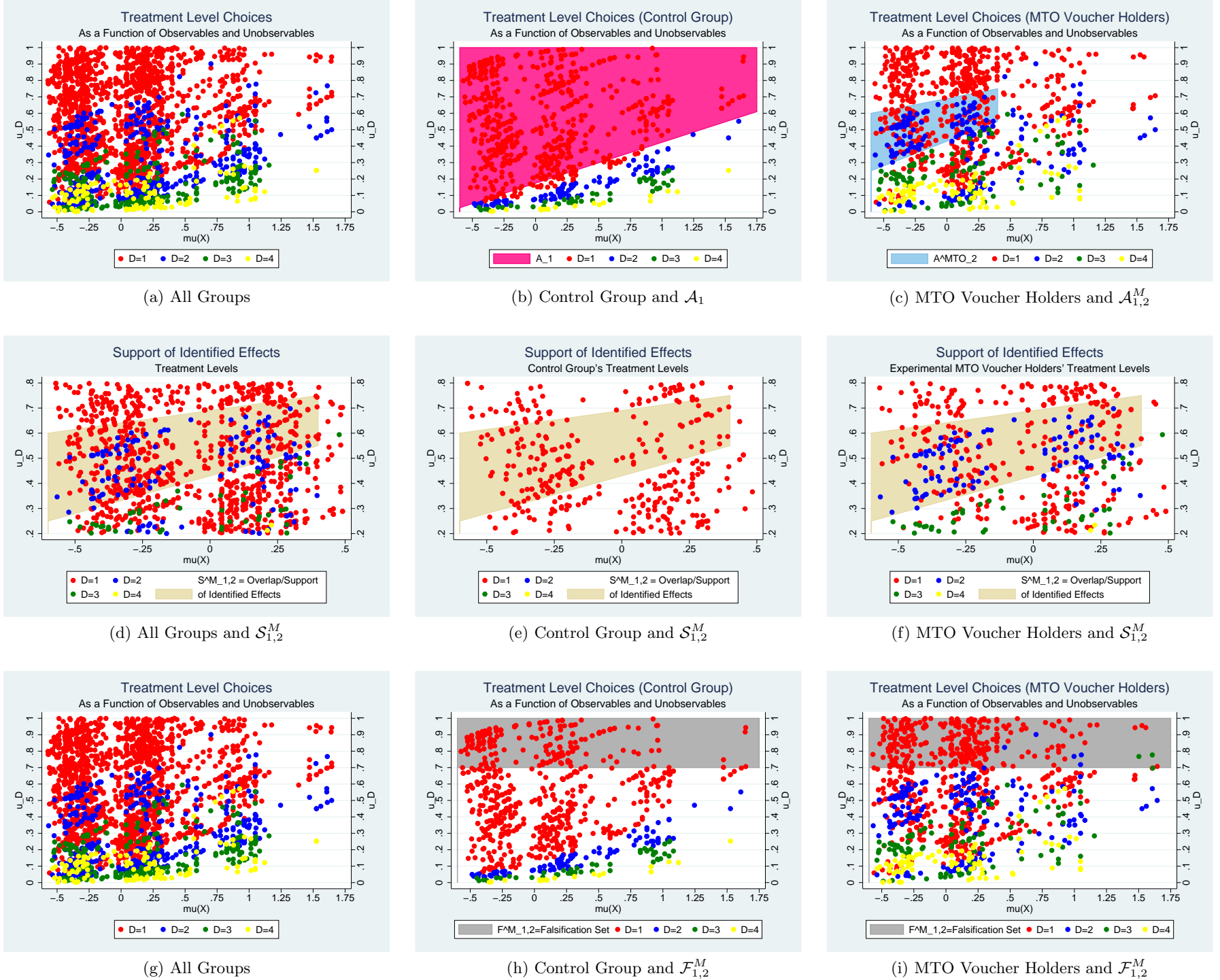


Figure 10: Selection into Treatment, Counterfactual Areas \mathcal{A}_1 and $\mathcal{A}_{1,2}^M$, Identification Support Set $\mathcal{S}_{1,2}^M$, and Falsification Set $\mathcal{F}_{1,2}^M$

Note: This Figure again shows how MTO adults sorted into the discrete quality levels $D_i \in \{1, 2, 3, 4\}$ as a function of their estimated choice model observed characteristics ($\hat{\mu}(X_i)$ - on the x -axis) and unobserved characteristics (\hat{U}_{D_i} - on the y -axis). We now also show the construction of the identification of the support set using its components \mathcal{A}_1 in (b) and $\mathcal{A}_{1,2}^M$ in (c). The identification support set $\mathcal{S}_{1,2}^M = \mathcal{A}_1 \cap \mathcal{A}_{1,2}^M$ is shown in (d), and then for the control group and MTO voucher holders separately in (e) and (f). Finally, the falsification set $\mathcal{F}_{1,2}^M$ is shown in Figures (g)-(i), which is the group of households identified by the choice model to counterfactually all have $D_i(Z_i^M = 0) = D_i(Z_i^M = 1) = 1$.

Table 6: Adult LATE Estimates

Outcome	$\Delta_{1,2}^{LATE}(Z^M)$	Control Mean in \mathcal{S}_1^M
Labor Market		
In Labor Force (%)	25.8** (18.3)	53.2
Employed (%)	31.2** (20.1)	41.7
Household Income (\$)	5,616* (3,914)	13,506
Earnings (\$)	1,970 (4,066)	7,642
Welfare Benefits		
Received TANF (%)	-40.0*** (19.2)	39.9
Health		
BMI (Raw)	-3.1 (2.8)	30.9

Note: $\Delta_{1,2}^{LATE}(Z^M)$ estimates pertain to individuals with observed and unobserved choice model components in $\mathcal{S}_1^M \equiv \{(\mu(X_i), U_{Di}) | \mu(X_i) \in [-0.6, 0.4], u_{Di} \in [0.43 + 0.30\mu(X_i), 0.68 + 0.15\mu(X_i)]\}$. Control means are also computed for the subsample in this region and outside of this region (both conditional on $D_i = 1$). Standard errors are computed using 200 bootstrap replications, with * denoting statistical significance at the (one-tailed) 10% level determined using either the 10th or 90th percentile of the 200 bootstrapped replications, and ** and *** denoting significance at the 5% and 1% levels, respectively.

5.3.4 Falsification Test: Outcomes for Non-Complier Households

To highlight the difference between our analysis and the program effect approach adopted in most of the literature on MTO (Ludwig et al. (2008), Ludwig et al. (2013), Chetty et al. (2016)), we now use Figure 10 to define a falsification set $\mathcal{F}_{1,2}^M$ for which households would remain in neighborhoods of quality $D = 1$ even if they were assigned an MTO voucher:

$$\mathcal{F}_{1,2}^M \equiv \left\{ (\mu(X_i), U_{Di}) \mid U_{Di} \in [0.70, 1.0] \right\}.$$

Households in the falsification set are determined not only by their values of $(\mu(X_i), U_{Di})$, but also by the general cost function $C(q)$ and the value of the cost reduction function $\gamma^M(q)$ at various levels of quality q .

Neighborhood selection for the subpopulations in $\mathcal{S}_{1,2}^M$ and $\mathcal{F}_{1,2}^M$ is shown in the CDFs in Figure 11. Figure 11a shows that there is considerable variation in the neighborhood quality selected by the control and MTO voucher holders in the identification support set $\mathcal{S}_{1,2}^M$: No households in the control group selected into $D = 2$, while 37 percent of MTO voucher holders did. Neighborhood selection for households in the falsification set $\mathcal{F}_{1,2}^M$ was quite different: No households in the control group selected into $D = 2$, and only 2 percent of MTO voucher holders selected into $D = 2$.

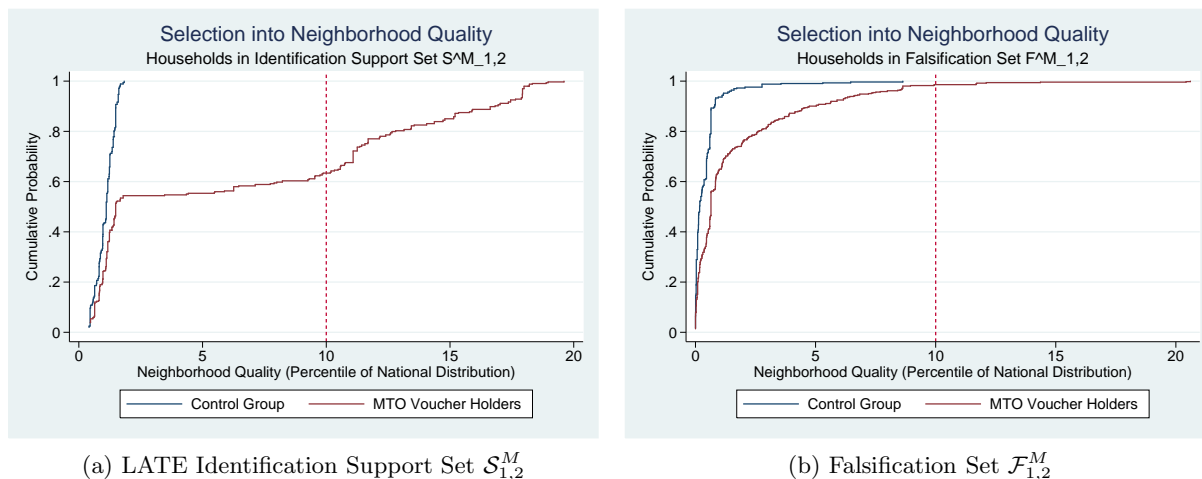


Figure 11: Selection into Neighborhood Quality for Various Subpopulations

Note: This Figure shows the distributions of $q_i(Z_i^M = 0)$ and $q_i(Z_i^M = 1)$ in (a) for the identification support set $\mathcal{S}_{1,2}^M$ and in (b) for the falsification set $\mathcal{F}_{1,2}^M$. We see that the model does seem to identify groups whose counterfactual neighborhood selection is quite different, and that selection into the continuous measure of neighborhood quality is consistent with the selection into the discrete measure of neighborhood quality shown earlier.

The effects of the MTO program are compared in Table 7 for households in the LATE identification support set $\mathcal{S}_{1,2}^M$ and for households in the falsification set $\mathcal{F}_{1,2}^M$. While receiving an MTO voucher resulted in large improvements to labor force participation rates for households with characteristics making them likely to move to a higher quality neighborhood when offered an MTO voucher (ie, those in $\mathcal{S}_{1,2}^M$), there was no effect on labor force participation for households receiving the MTO voucher who did not move to a higher quality neighborhood. Employment actually went down for those who did not move to higher quality neighborhoods, perhaps due to the disruptiveness of moving without the benefits of moving closer to jobs (Weinberg (2000)). And while welfare (TANF) receipt and BMI decreased for voucher recipients who did not move to higher quality neighborhoods, this effect was between multiple times and an order of magnitude larger for those who did move to a higher quality neighborhood.

This falsification test helps to illustrate that the effects of the MTO program are not interchangeable with effects from neighborhood quality. A list of assumptions must be made before translating effects of variation in MTO voucher assignment into effects of variation in neighbor-

hood quality. Our assumptions have been stated explicitly in Sections 2-5; most assumptions in the MTO literature have been made implicitly (Aliprantis (2015a)).

Table 7: Adult Program Effects Estimates by Neighborhood Selection Groups

	Falsification Set: $(\mu(X_i), U_{Di}) \in \mathcal{F}_{1,2}^M$ (No Change in Nbd Quality due to MTO)			Identification Support Set: $(\mu(X_i), U_{Di}) \in \mathcal{S}_{1,2}^M$ (Some Improvement in Nbd Quality due to MTO)		
	$E[Y Z^M = 1]$	$E[Y Z^M = 0]$	$E[Y Z^M = 1] - E[Y Z^M = 0]$	$E[Y Z^M = 1]$	$E[Y Z^M = 0]$	$E[Y Z^M = 1] - E[Y Z^M = 0]$
Neighborhood Selection						
Nbd Quality ($D \in \{1, 2, 3, 4, 5\}$)	1.02 (0.07)	1.00 (0.00)	0.02 (0.07)	1.37 (0.10)	1.00 (0.09)	0.37 (0.13)
Nbd Quality ($q \in [0, 50]$)	1.7 (0.8)	0.4 (0.1)	1.2 (0.8)	6.4 (1.0)	1.1 (0.9)	5.3 (1.3)
Labor Market						
In Labor Force (%)	63.6 (3.9)	63.6 (5.4)	0.0 (6.4)	63.0 (3.2)	53.2 (3.8)	9.8 (4.7)
Employed (%)	47.1 (4.2)	53.6 (5.4)	-6.5 (6.8)	53.5 (3.3)	41.7 (3.9)	11.8 (5.0)
Household Income (\$)	14,252 (924)	14,134 (998)	119 (1,366)	15,629 (847)	13,506 (883)	2,123 (1,175)
Earnings (\$)	7,583 (914)	8,554 (992)	-971 (1,375)	8,364 (611)	7,642 (767)	722 (917)
Welfare Benefits						
Received TANF (%)	32.2 (3.7)	33.7 (5.0)	-1.5 (6.6)	24.9 (3.0)	39.9 (3.4)	-15.0 (4.6)
Health						
BMI (Raw)	30.0 (0.5)	30.4 (0.8)	-0.3 (1.0)	29.7 (0.5)	30.9 (0.5)	-1.2 (0.7)

Note: The first three columns of this table report the effects of receiving an experimental MTO voucher for households predicted by the estimated choice model to reside in a low-quality neighborhood even when receiving an MTO voucher. The last three columns report the effects of receiving an experimental MTO voucher for households predicted by the estimated choice model to potentially move to a higher quality neighborhood when receiving an MTO voucher.

6 Discussion

Aliprantis (2015a) clarifies the obstacles to learning about neighborhood effects from program effect estimates, given that compliance with random assignment often led to little or no improvements in neighborhood quality. Drawing on the work of Heckman et al. (2006), we specified a joint model of neighborhood quality choice and outcomes, allowing unobserved heterogeneity to influence moving decisions with and without voucher assignment. Since we observe the precise level of neighborhood quality selected by households, the structure of the neighborhood choice model allows us to estimate observed and unobserved components of choice in the absence of a voucher for all program participants. With households thus classified, we are able to obtain LATE estimates for the most common neighborhood quality transition induced by MTO: a move from the first to the second decile of neighborhood quality.

The fact that LATE estimates are large for labor market outcomes, despite the relatively small improvement in neighborhood quality, strongly suggests that MTO does provide evidence of positive neighborhood effects. However, these benefits are realized by a small percentage of voucher holders. Not only was the MTO voucher take-up rate low at 43 percent, but an upward move in quality was even less likely. The LATEs we estimated pertain to about 9 percent of MTO volunteers. And MTO volunteers make up only about a quarter of MTO-eligible households (Chyn (2015))).

We interpret our estimates' lack of generality as evidence that policy making should be a dynamic process that is responsive to the available evidence (Manski (2013b)). Despite the fact that households were more likely to move with Section 8 vouchers than MTO vouchers, this does not imply that Section 8 vouchers are preferable to MTO vouchers. Changes in neighborhood quality were much smaller for Section 8 movers than for MTO movers, and variation by site was large. Since only about a quarter of eligible households are currently able to obtain a Section 8 housing voucher (Sard and Fischer (2012)), an area for future research is understanding what changes in voucher policy might optimize the extent to which households are able to realize positive neighborhood effects through the subsidy, and which of these policies might be feasible to implement (McClure (2010), Collinson and Ganong (2013)).

A final consideration when interpreting our results is that programs and neighborhood changes differentially impacting the treatment and control groups will interfere with the ability to identify neighborhood effects using voucher assignment as an instrument. Ideally, the voucher assignment would induce a change in the cost of moving holding all else equal. However, households may have responded to their group assignment, and baseline neighborhood conditions may well have changed during the multiple-year period between the decision to move and the time of the interim evaluation when outcomes were measured. For example, households assigned to the control group might have responded by applying for Section 8 vouchers on their own outside of the MTO program (Orr et al. (2003)). And according to de Souza Briggs et al. (2010), during the implementation of MTO Jobs-Plus saturated public housing developments with state-of-the-art employment, training, and child care services, while providing rent incentives to encourage employment. The US also enacted major welfare reform legislation in August 1996, precisely while MTO vouchers were being assigned

(Blank (2002)). That these and other factors are not explicitly modeled in our analysis points to the methodological limitations of conducting controlled, and not just randomized, experiments in social settings (Deaton (2010), Aliprantis (2015b)).

7 Conclusion

Because households endogenously sort into neighborhoods, identifying causal effects of neighborhood environments has proven to be a substantial challenge. Researchers have sought to identify neighborhood effects using the exogenous variation in neighborhoods induced by housing mobility programs. The Moving to Opportunity (MTO) housing mobility experiment gave households living in high-poverty neighborhoods in five US cities the ability to enter a lottery for housing vouchers to be used in low-poverty neighborhoods. In a tremendous disappointment, effects from MTO were very small, and this lack of effects has been interpreted as evidence against the theory that neighborhood characteristics influence individuals' outcomes.

Aliprantis (2015a) clarifies that program effect estimates are less informative about neighborhood effects than currently appreciated in the literature, given that compliance with random assignment often led to little or no improvements in neighborhood quality. This paper explored an alternative to learning about neighborhood effects from ITT and TOT effects of the MTO program. We proposed and implemented a new strategy for identifying transition-specific LATEs of neighborhood quality using the random variation in quality induced by the program, based on a model of households' selection into neighborhood quality.

We found that moving to a higher quality neighborhood had large, positive effects on employment, labor force participation, household income, and welfare receipt. Although effects on individual earnings and body mass index (BMI) were not statistically significant, these effects were also estimated to be large and positive. Due to the limited changes in neighborhood quality induced by MTO, these LATE estimates pertain to about nine percent of our estimation sample. We found no evidence from MTO against the theory that increasing neighborhood quality improves adult outcomes.

References

- Aliprantis, D. (2012). Redshirting, compulsory schooling laws, and educational attainment. *Journal of Educational and Behavioral Statistics* 37(2), 316–338.
- Aliprantis, D. (2014). Human capital in the inner city. *Federal Reserve Bank of Cleveland Working Paper 13-02r*.
- Aliprantis, D. (2015a). Assessing the evidence on neighborhood effects from Moving to Opportunity. *Federal Reserve Bank of Cleveland Working Paper 15-06*.

- Aliprantis, D. (2015b). Covariates and causal effects: The problem of context. *Federal Reserve Bank of Cleveland Working Paper 13-10R*.
- Altonji, J. and R. Mansfield (2014). Estimating school and neighborhood effects using sorting on observables to control for sorting on unobservables. *Mimeo., Cornell University*.
- Anderson, E. (1999). *Code of the Street: Decency, Violence, and the Moral Life of the Inner City*. W. W. Norton and Company.
- Angrist, J. D. and G. W. Imbens (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association* 90(430), 431–442.
- Bayer, P., R. McMillan, A. Murphy, and C. Timmins (2011). A dynamic model of demand for houses and neighborhoods. *NBER Working Paper 17250*.
- Bayer, P. J. and S. L. Ross (2009). Identifying Individual and Group Effects in the Presence of Sorting: A Neighborhood Effects Application. *University of Connecticut Working Paper 2006-13R*.
- Blank, R. M. (2002). Evaluating welfare reform in the United States. *Journal of Economic Literature* 40(4), 1105–1166.
- Brock, W. and S. Durlauf (2007). Identification of binary choice models with social interactions. *Journal of Econometrics* 140(1), 52–75.
- Chetty, R., N. Hendren, and L. F. Katz (2016, April). The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment. *American Economic Review* 106(4), 855–902.
- Chyn, E. (2015). Moved to Opportunity: The long-run effect of public housing demolition on labor market outcomes of children. *Mimeo., University of Michigan*.
- 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.
- Collinson, R. and P. Ganong (2013). Incidence and price discrimination: Evidence from housing vouchers. *Harvard University Joint Center for Housing Studies Working Paper W13-7*.
- Dale, S. B. and A. B. Krueger (2002). Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *The Quarterly Journal of Economics* 117(4), 1491–1527.
- de Souza Briggs, X., S. J. Popkin, and J. Goering (2010). *Moving to Opportunity: The Story of an American Experiment to Fight Ghetto Poverty*. Oxford University Press.

- Deaton, A. (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature* 48(2), 424–455.
- 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.
- DeLuca, S. and P. Rosenblatt (2010). Does moving to better neighborhoods lead to better schooling opportunities? Parental school choice in an experimental housing voucher program. *Teachers College Record* 112(5), 1443–1491.
- Fletcher, J. M. and S. L. Ross (2012). Estimating the effects of friendship networks on health behaviors of adolescents. *NBER Working Paper 18253*.
- Florens, J.-P., J. J. Heckman, C. Meghir, and E. Vytlacil (2008). Identification of treatment effects using control functions in models with continuous, endogenous treatment and heterogeneous effects. *Econometrica* 76(5), 1191–1206.
- Fu, S. and S. L. Ross (2013). Wage premia in employment clusters: How important is worker heterogeneity? *Journal of Labor Economics* 31(2), 271–304.
- Galiani, S., A. Murphy, and J. Pantano (2015). Estimating neighborhood choice models: Lessons from a housing assistance experiment. *The American Economic Review* 105(11), 3385–3415.
- 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).
- Greene, W., M. N. Harris, B. Hollingsworth, and P. Maitra (2008). A bivariate latent class correlated generalized ordered probit model with an application to modeling observed obesity levels. *Mimeo., New York University*.
- Heckman, J. J. and S. Urzúa (2010). Comparing IV with structural models: What simple IV can and cannot identify. *Journal of Econometrics* 156(1), 27–37.
- Heckman, J. J., S. Urzúa, and E. Vytlacil (2006). Understanding Instrumental Variables in models with essential heterogeneity. *The Review of Economics and Statistics* 88(3), 389–432.
- Heckman, J. J. and E. Vytlacil (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.
- 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.
- Kim, K. and A. Petrin (2013). A new control function approach for non-parametric regressions with endogenous variables. *NBER Working Paper 16679*.
- Kling, J. R., J. B. Liebman, and L. F. Katz (2005). Bullets don't got no name: Consequences of fear in the ghetto. In T. S. Weisner (Ed.), *Discovering Successful Pathways in Children's Development: Mixed Methods in the Study of Childhood and Family Life*, pp. 243–281. University of Chicago Press.
- Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental analysis of neighborhood effects. *Econometrica* 75(1), 83–119.
- 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.
- Lochner, L. and E. Moretti (2015). Estimating and testing models with many treatment levels and limited instruments. *The Review of Economics and Statistics* 97(2), 387–397.
- Ludwig, J., G. J. Duncan, L. A. Gennetian, L. F. Katz, R. C. Kessler, J. R. Kling, and L. Sanbonmatsu (2013). Long-term neighborhood effects on low-income families: Evidence from Moving to Opportunity. *American Economic Review* 103(3), 226–231.
- Ludwig, J., G. J. Duncan, and J. C. Pinkston (2005). Housing mobility programs and economic self-sufficiency: Evidence from a randomized experiment. *Journal of Public Economics* 89(1), 131–156.
- 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.
- Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies* 60(3), pp. 531–542.
- Manski, C. F. (2013a). Identification of treatment response with social interactions. *The Econometrics Journal* 16(1), S1–S23.
- Manski, C. F. (2013b). *Public Policy in an Uncertain World: Analysis and Decisions*. Harvard University Press.

- Massey, D. and N. Denton (1993). *American Apartheid: Segregation and the Making of the Underclass*. Cambridge: Harvard University Press.
- McClure, K. (2010). The prospects for guiding housing choice voucher households to high opportunity neighborhoods. *Cityscape* 12(3), 101–122.
- 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>.
- Nekipelov, D. (2011). Identification and semiparametric efficiency in a model with a multi-valued discrete regressor. *Mimeo., UC Berkeley*.
- Oreopoulos, P. (2003). The long-run consequences of living in a poor neighborhood. *The Quarterly Journal of Economics* 118(4), pp. 1533–1575.
- Orr, L. L., J. D. Feins, R. Jacob, E. Beecroft, L. Sanbonmatsu, L. F. Katz, J. B. Liebman, and J. R. Kling (2003). *Moving to Opportunity: Interim Impacts Evaluation*. Washington, DC: US Department of Housing and Urban Development, Office of Policy Development and Research.
- Pinto, R. (2014). Selection bias in a controlled experiment: The case of Moving to Opportunity. *Mimeo., University of Chicago*.
- 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.
- 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. (2012). *Great American City: Chicago and the Enduring Neighborhood Effect*. The University of Chicago Press.
- 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.
- Sard, B. and W. Fischer (2012, July). Renter’s tax credit would promote equity and advance balanced housing policy. *Center on Budget and Policy Priorities Report*.

- 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.
- Weinberg, B. A. (2000). Black residential centralization and the spatial mismatch hypothesis. *Journal of Urban Economics* 48(1), 110–134.
- Wilson, W. J. (1987). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. University of Chicago.