Evidence on the Production of Cognitive Achievement from Moving to Opportunity

Dionissi Aliprantis*

June 22, 2017

Abstract: This paper performs a subgroup analysis on the effect of receiving a Moving to Opportunity (MTO) housing voucher on test scores. I find evidence of heterogeneity by number of children in the household in Boston, gender in Chicago, and race/ethnicity in Los Angeles. To study the mechanisms driving voucher effect heterogeneity, I develop a generalized Rubin Causal Model and propose an estimator to identify transition-specific Local Average Treatment Effects (LATEs) of school and neighborhood quality. Although I cannot identify such LATEs with the MTO data, the analysis demonstrates that membership in a specific demographic group is more predictive of voucher effects than is the group's average change in school or neighborhood quality. I discuss some possible explanations.

Keywords: Two-Dimensional Treatment, Rubin Causal Model, School Effect, Neighborhood Effect, Local Average Treatment Effect, Education Production Function (EPF), Moving to Opportunity
JEL Classification Numbers: C31, C36, C50, D04, I20, I38, R23

*Federal Reserve Bank of Cleveland, PO Box 6387, Cleveland, OH 44114, *dionissi.aliprantis@clev.frb.org*, +1(216)579-3021

I thank Francisca G.-C. Richter, Bruce Fallick, Peter Hinrichs, and Hal Martin for helpful comments, Samantha Burg at the US Department of Education for help with the NAEP data, as well as Doctor Ashe and Arun Mathur at ICPSR and Stephanie Tulley at the Cleveland Fed for help with the MTO data. The opinions expressed are those of the author and do not represent views of the Federal Reserve Bank of Cleveland or the Board of Governors of the Federal Reserve System.

1 Introduction

Receiving a Moving to Opportunity (MTO) housing voucher had no average effect on adult wages (Kling et al. (2007), Ludwig et al. (2013)), but recent studies have found positive voucher effects for specific subpopulations. Aliprantis and Richter (2016) and Pinto (2015) find large effects on economic self-sufficiency when focusing on those adults who would be induced by the experiment to improve neighborhood quality. Chetty et al. (2016) find voucher effects on subsequent wages and college attendance when focusing on children who were less than 13 when assigned a voucher.

This paper conducts a subgroup analysis of voucher effects that is guided by these recent results on MTO, which suggest that some subpopulation of children could have experienced voucher effects on test scores as a mediator of long-term outcomes. Estimating Intent-to-Treat (ITT) voucher effects by subgroups, I do find heterogeneous effects by the number of children in the household in Boston; gender in Chicago; and race/ethnicity in Los Angeles. I build on the subgroup analysis in Sanbonmatsu et al. (2006) by using new measures of school and neighborhood quality, and by addressing two of the major obstacles raised in their analysis: measurement error in test scores and missing data on school quality.

The paper then attempts to leverage the experimental design of MTO to understand the mechanisms driving this voucher effect heterogeneity. The experimental feature of MTO is that households were randomly assigned housing vouchers with restrictions to move to low-poverty neighborhoods. The randomized housing voucher is an instrument for school and neighborhood quality.

To achieve identification with the MTO instrument, I specify a joint model of school and neighborhood selection, together with a model of potential outcomes. Because school and neighborhood quality are likely to be distinct inputs into the production of cognitive achievement, we would like to understand their complementarity. What are their effects independently of one another, and how do they combine to impact achievement?

To accommodate school and neighborhood environments as distinct factors of production, I specify a selection model in which households choose both school and neighborhood quality. I also assume potential outcomes depend on school and neighborhood quality in a discrete way that is ordered into multiple levels in each dimension.

I show through the model that a simple estimator requiring only a discrete instrument like the MTO voucher can identify transition-specific Local Average Treatment Effects (LATEs) of changing school quality, neighborhood quality, or both. The goal of the model is to isolate those subgroups that were induced to make specific changes in school and neighborhood quality as a result of the program. A similar approach was taken in Aliprantis and Richter (2016), but must be modified here because the model in this paper works in two dimensions rather than one.

I cannot identify the LATEs of interest with the MTO data. Looking at the changes in school and neighborhood quality induced by the MTO voucher, I find that isolating those children making transition-specific moves would shrink the sample size prohibitively. The small sample problem is driven not only by the strength of the MTO instrument, but also by a limitation of the MTO data. Two of the five MTO sites cannot be used in the analysis because they do not include measures of

school quality in terms of standardized test scores.

Despite this negative identification result, we are able to learn from the analysis: A demographic group's average change in school and neighborhood quality is less predictive of voucher effects than is membership in the demographic group itself. There are several possible explanations for why voucher effects are more heterogeneous by demographic characteristics than by changes in school or neighborhood quality.

A first possibility is that changes in average school and neighborhood quality are misleading. Interpreting voucher effects in terms of school and neighborhood effects is difficult without first estimating a selection model capable of predicting, or at least restricting, individual-level counterfactual responses to the instrument. This point is well-appreciated in the literature evaluating the effects of MTO on youth outcomes (Gennetian et al. (2012), Sanbonmatsu et al. (2006)), but somewhat controversial in the literature on adult outcomes (Aliprantis (2017)).

Consider a demographic group that experienced an increase in average school and neighborhood quality when receiving an MTO voucher; their voucher effects could have resulted from many types of moves. The group's changes in average quality could have been driven by a subgroup that increased both school and neighborhood quality. However, the changes in quality could also have resulted from one subgroup greatly increasing school quality while slightly decreasing neighborhood quality, and a second subgroup slightly decreasing school quality while greatly increasing neighborhood quality.[1] We cannot interpret voucher effects without being able to distinguish between these scenarios.

A second possibility is that my measures of school and neighborhood quality do not measure the variables that are factors of production for cognitive achievement. Exposure to violence (Kling et al. (2005), Aliprantis (2016)) and activities outside of school (Zuberi (2010)) both matter, but are omitted from my measures of quality. Value-added measures of schools and teachers (Rockoff et al. (2014)) or neighborhoods (Davis et al. (2017)) might be more appropriate than my measures. And the peer effects experienced in the classroom might not be well-measured by the average test score in a school (Tincani (2015), Fruehwirth (2014)).

A third possibility is that moves along the margins of quality experienced in MTO do not matter. This would be the conclusion if selection and measurement were not issues, so that the voucher effects did capture the effects of moving to higher quality schools and neighborhoods. This could be because the change in environment along these margins is not large; the average long-term change in school percentile was 3 percentile points (Gennetian et al. (2012)). It could also be that responses along these margins are not large.

A fourth possibility is that reading (or non-cognitive) skills are the mediator improving long-term wages, and I have only examined effects on math skills. This is again due to a limitation of the MTO data, in that interviewer judgment seems to have influenced test scores (See Appendix 1 of Sanbonmatsu et al. (2006)). A fifth possibility is that even the math test scores used in the

[1] This would be a violation of monotonicity in each dimension. For theoretical discussions of monotonicity in one dimension see Angrist and Imbens (1995) and Heckman et al. (2006), and for an empirical example see Aliprantis (2012) and Barua and Lang (2016).

analysis suffer from this biased measurement. Finally, it is possible that my static model does not give adequate attention to the dynamics of the data generating process.

The remainder of the paper is structured as follows: Section 2 describes the MTO experiment, and Section 3 describes the data used in the analysis. Empirical results on voucher effects are presented in Section 4. Section 5 presents a joint model of selection and potential outcomes and defines causal effects of interest, with Section 6 presenting an estimator capable of identifying causal effects of school and neighborhood quality. Empirical results on school and neighborhood effects are presented in Section 7, and Section 8 concludes.

2 Moving to Opportunity (MTO)

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)).[2]

MTO was designed to replicate these beneficial effects, offering housing vouchers to eligible households between September 1994 and July 1998 in Baltimore, Boston, Chicago, Los Angeles, and New York (Goering (2003)). Households were eligible to participate in MTO if they were low-income, had at least one child under 18, were residing in either public housing or Section 8 project-based housing located in a census tract with a poverty rate of at least 40%, were current in their rent payment, and all 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

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

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.[3]

3 Data

The sources of data used in the analysis are the MTO Interim Evaluation restricted-access data set, national data on the distribution of National Assessment of Educational Progress (NAEP) test scores, and tract-level data from the 2000 decennial US Census obtained from the National Historical Geographic Information System (NHGIS, Minnesota Population Center (2004)).

I use math test scores as the outcome of interest, average math test score in a school as the measure of school quality, and a combination of census tract characteristics as the measure of neighborhood quality.

I restrict the analysis to Boston, Chicago, and Los Angeles. As noted in Sanbonmatsu et al. (2006) (See Table 2 and Figure 1), school exam scores were not available for older children in Baltimore and New York City. It is therefore not possible to rank these children's school quality using performance on a standardized test.

3.1 Outcomes

I focus the analysis on math test scores because the interviewers administering reading tests appear to have graded the tests differently. Appendix 1 of Sanbonmatsu et al. (2006) discusses two sources of bias in the test score measurements of students participating in MTO due to the fact that it was largely a verbal assessment. First, students may have understood some interviewers better than others. Second, interviewers may have systematically understood some students better than others. To overcome these problems, I use the Calculation (CA) subsection of the math Woodcock-Johnson Revised (WJR) test scores. For the CA WJR subsection the interviewers did not read any of the questions aloud.[4]

3.2 School Quality

School quality is based on a national measure that is comparable across MTO sites. For each child's current school, the MTO Interim Evaluation contains a variable reporting where the school's average test score ranks in the state's distribution (individual-level) of test scores.[5] I create a measure of school quality by mapping each state-level percentile ranking into a percentile ranking

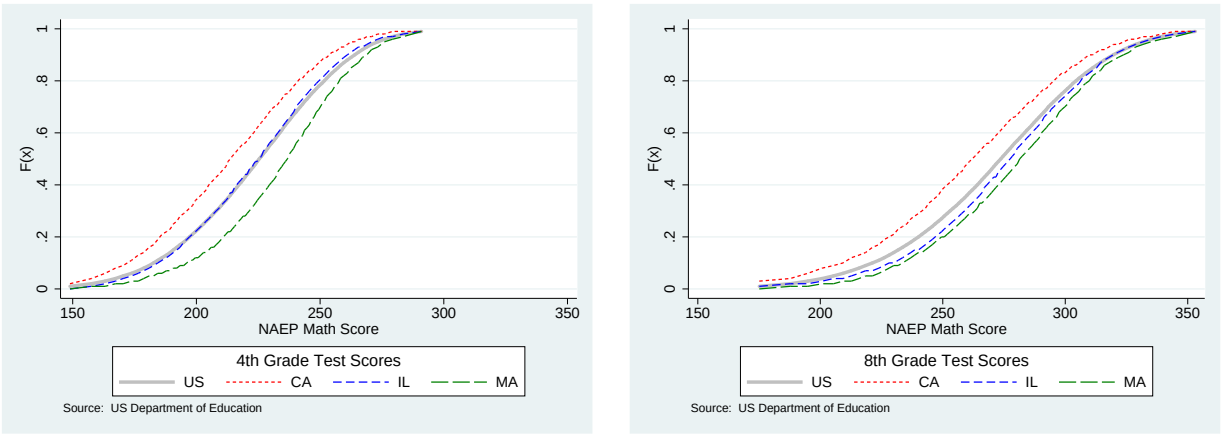
[3] Section 8 vouchers pay part of a tenant's private market rent. Project-based assistance gives the option of a reduced-rent unit tied to a specific structure.

[4] I also run the analysis using math test scores that are an average of the CA subsection and the Applied Problems (AP) subsection. The results are qualitatively similar, but I focus on the CA outcomes here because all of the AP questions were read aloud by the interviewer.

[5] Average ranking over the years 1999, 2000, and 2001 for the school the child was attending at the time of the interim evaluation survey in 2002.

of the national distribution (individual-level) of test scores. I do this using data provided by the US Department of Education listing each percentile for the nation, Massachusetts, Illinois, and California. These percentiles were available for the 4th and 8th grades in the year 2000.[6]

How much does this matter? Figure 1b illustrates the issue. Consider two children who attended middle schools whose average raw math test score on the NAEP was a 224. If the first child attended a school in Boston (Massachusetts), her school would be at the 7th percentile in the state. However, if the second child attended a school in Los Angeles (California), her school would be at the 18th percentile of her state. To ensure that we are measuring movements across similar margins of school quality in all three of the MTO states under investigation, I map both of these schools into the national distribution. According to this measure, both girls are attending a school at the 11th percentile of school quality. Figure 1a below shows that there are even larger discrepancies across states in elementary school.



(a) 4th Grade

(b) 8th Grade

Figure 1: State and National Distributions of NAEP Test Scores (2000)

3.3 Neighborhood Quality

I use the neighborhood quality measure from Aliprantis (2017) and Aliprantis and Richter (2016) that is constructed with 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. A census tract's quality is the percentile of the first principal component of these variables.

[6] I use 4th grade rankings for children aged 5–11 at the time of the interim evaluation and 8th grade rankings for children aged 12-19 in the interim evaluation.

3.4 Demographic Characteristics

As discrete demographic variables I initially considered the following variables:

$$\begin{aligned}
\text{Site} &\in \{\text{Boston , Chicago , Los Angeles }\} \\
\text{Mother's Highest Degree} &\in \{\text{Dropout}, \geq \text{GED}\} \\
\text{Number of Children in Household} &\in \{\leq 2, \geq 3\} \\
\text{Black} &\in \{0, 1\} \\
\text{Female} &\in \{0, 1\} \\
\text{Baseline Neighborhood Quality} &\in \{< \text{Site Median} , \geq \text{Site Median}\}
\end{aligned}$$

I include site as a demographic variable because the selection patterns were so different by site (Section 7 shows these selection patterns in detail.). I then restrict the demographic variables to three per site to maintain sample sizes. Finally, I choose the specific characteristics at each site to maximize the differences of treatments between the Control and Experimental MTO groups. The characteristics used in the analysis for each site are

Boston:	Black,	Female,	Number of Children in Household;
Chicago:		Female, Mother's Highest Degree,	Number of Children in Household;
Los Angeles:	Black,	Mother's Highest Degree,	Number of Children in Household.

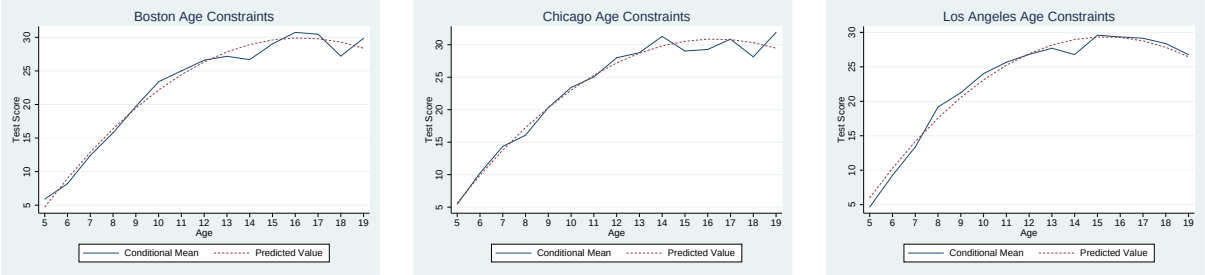
4 Empirical Results on Voucher Effects

A post hoc subgroup analysis is subject to the problem of multiple comparisons, and so may easily generate false positives. This is one reason for subsequently trying to estimate causal effects of school and neighborhood quality of the type we would expect to find before looking at the MTO data. To be clear, the results here are suggestive of possible mediating mechanisms, and are presented as a hypothesis-generating exercise. Conditional on further access to the MTO Interim Evaluation restricted-access data, a more formal subgroup analysis will be conducted along the lines of Imai and Ratkovic (2013), Tian et al. (2014), or Wager and Athey (2017).

To accommodate age effects in small subsamples, I first estimate site-level regressions of test scores on a quadratic function of age:

$$T_i = \beta_0 + \beta_1 age_i + \beta_2 age_i^2 + \epsilon_i. \tag{1}$$

These coefficients are used as constraints in subsequent regressions. The fit of these regressions is shown below in Figure 2:



(a) Boston (b) Chicago (c) Los Angeles

Figure 2: Average Test Scores as a Quadratic Function of Age

I study voucher effect heterogeneity as the $\beta_{.Z}$ coefficients on site- and demographic-group-specific, constrained least-squares regressions

$$\begin{aligned} q_{Si} &= \beta_{S0} + \beta_{SZ} Z_i^M + \epsilon_{Si} \\ q_{Ni} &= \beta_{N0} + \beta_{NZ} Z_i^M + \epsilon_{Ni} \\ T_i &= \beta_{T0} + \beta_{T1} age_i + \beta_{T2} age_i^2 + \beta_{TZ} Z_i^M + \epsilon_{Ti} \end{aligned} \tag{2}$$

where q_{Si} is school quality, q_{Ni} is neighborhood quality, T_i is WJR Math Calculation subtest score, $Z_i^M = 1$ for MTO voucher holders and $Z_i^M = 0$ for the control group, $X_i = k$ is an observed characteristic, and the coefficients on age and age^2 are constrained in Equation 2 to their estimated values from the site-specific regressions in Equation 1:

$$\beta_{T1} = \widehat{\beta}_1 \qquad \text{(Constraint 1)}$$

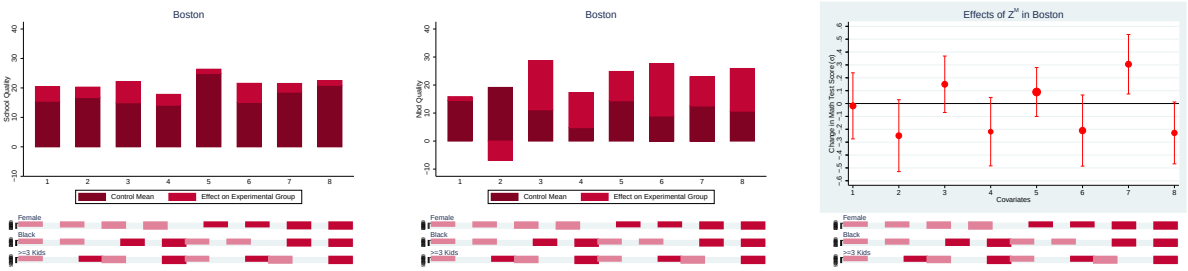
$$\beta_{T2} = \widehat{\beta}_2 \qquad \text{(Constraint 2)}$$

Figures 3-5 show the experimental voucher effects β_{SZ}, β_{NZ}, and β_{TZ} by site and demographic characteristics. The bottom panels record the demographic group of interest. For example, in Boston group 1 (G1) comprises children whose head of household has not attained a high school diploma or GED, has two kids or less in their household, and who is not black. Group 2 comprises children whose head of household has not attained a high school diploma or GED, has two kids or less in their household, but who is black.

In Boston, everyone with more than three kids in the household did worse from getting a voucher, regardless of gender or race/ethnicity. Those with two kids or less all did well, with the exception of G1, which is not a surprise since school quality went down and neighborhood quality remained the same. Two stylized facts from Chyn et al. (2017) that might be useful in interpreting these results is that Section 8 lease-up is elevated for households with three or more children and for children with poor recent academic performance.

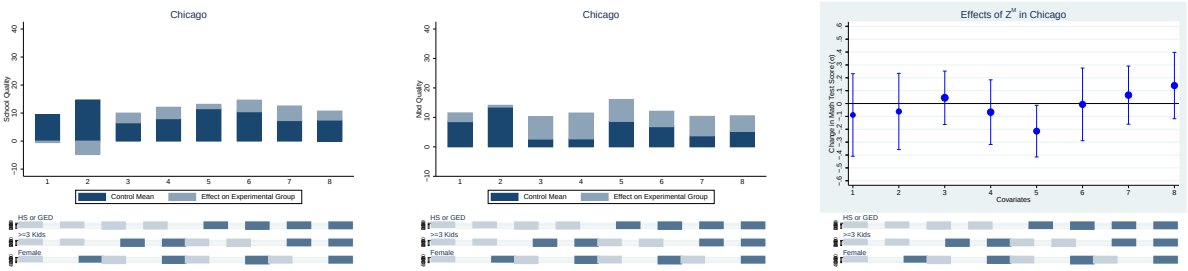
In Chicago, one might interpret girls as having benefited from receiving an MTO voucher even when experiencing very small improvements in school quality, while boys were disadvantaged from receiving an MTO voucher.

In Los Angeles, Hispanic children started out in better schools, but experienced no improvements in school quality. In contrast, black children did experience improvements in school quality. Hispanic children benefited from getting an MTO voucher, while black children fared worse as a result. This does not seem to be driven by changes in neighborhood quality.



(a) School Quality (b) Neighborhood Quality (c) Effects

Figure 3: Boston by Demographic Group



(a) School Quality (b) Neighborhood Quality (c) Effects

Figure 4: Chicago by Demographic Group



(a) School Quality (b) Neighborhood Quality (c) Effects

Figure 5: Los Angeles by Demographic Group

Since the effects on school and neighborhood quality are not huge, we might be worried that any effect heterogeneity is driven by sampling error. To investigate this issue (slightly) more formally, I estimate constrained regressions analogous to Equation 2 for test scores that include an interaction term:

$$T_i = \beta_0 + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 Z_i^M + \beta_4 Z_i^M \times \mathbf{1}\{X_i\} + \epsilon_{Ti} \tag{3}$$

where $\mathbf{1}\{X_i\}$ is an indicator for a given set of observed characteristics (number of kids at home, gender, or race/ethnicity).

Assuming that the visual inspection of Figures 3c, 4c, and 5c amounts to performing independent hypothesis tests for 3 binary covariates at each of the 3 sites, the Bonferroni correction to control for the family-wise error rate (the probability that one or more Type I errors will occur) at the 5 percent level will require comparing each p-value against a critical value of $p = 0.05/9 = 0.00\bar{5}$.

Table 1 shows that in regressions that adjust for age, each of the three types of voucher effect heterogeneity noted earlier are statistically significant: by the number of children in the household in Boston, by gender in Chicago, and race/ethnicity in Los Angeles. Table 2 in Appendix C shows that only the race/ethnicity heterogeneity survives in regressions that do not adjust for age. However, the p-values are of a magnitude to suggest that these coefficients represent true voucher effect heterogeneity, and not sampling variation. This is precisely the point of such an exploratory analysis: To focus our attention on possible explanations for the aggregate results we have observed.

Table 1: Effect Heterogeneity

Observed Characteristic	σ	t-statistic	P-Value
≥3 Kids in HH in Boston			
Coefficient on Z^M	0.07	1.20	0.231
	(0.06)		
Coefficient on Interaction Term	–0.27	–3.86	0.000
	(0.07)		
Female in Chicago			
Coefficient on Z^M	–0.15	–2.29	0.022
	(0.06)		
Coefficient on Interaction Term	0.31	4.50	0.000
	(0.07)		
Black in Los Angeles			
Coefficient on Z^M	0.06	0.89	0.373
	(0.07)		
Coefficient on Interaction Term	–0.31	–4.01	0.000
	(0.08)		

Note: Site-specific constrained least-squares regressions of

$$T_i = \beta_0 + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 Z_i^M + \beta_4 Z_i^M \times \mathbf{1}\{\boldsymbol{X_i} = k\} + \epsilon_{Ti}$$

where T_i is WJR Math Calculation subtest score, $Z_i^M = 1$ for MTO voucher holders and $Z_i^M = 0$ for the control group, $\boldsymbol{X_i} = k$ is the observed characteristic specified in the table, and β_1, and β_2 are constrained to their values from the site-specific regressions

$$T_i = \beta_0 + \beta_1 age_i + \beta_2 age_i^2 + \epsilon_i.$$

All regressions are weighted, P-Value is from a two-tailed test.

5 A Joint Model of School Choice, Neighborhood Choice, and Potential Outcomes

To further understand the mechanisms driving the effects of MTO vouchers, I now define causal effects of school and neighborhood quality. Suppose that both school and neighborhoods can be linearly ordered in terms of quality. Let $q_{Si} \in [0, 100]$ denote the percentile of quality of the school attended by child i, and let $q_{Ni} \in [0, 100]$ denote the percentile of quality of the neighborhood in which she resides. In the model households jointly choose their school and neighborhood, and this choice is a function of the child's observed ($\boldsymbol{X_i}$) and unobserved ($\boldsymbol{V_i}$) characteristics, as well as an instrument Z_i, according to the latent indexes

A1: $\quad (q^*_{Si}, q^*_{Ni}) = \boldsymbol{\mu_{S,N}}(\boldsymbol{X_i}, Z_i) - \boldsymbol{V_i}$

where $\boldsymbol{\mu_{S,N}}(\boldsymbol{X_i}) \in \mathcal{M} \equiv \mathbb{R}^2$ and $\boldsymbol{V_i} \equiv (V_{Si}, V_{Ni}) \in \mathcal{V} \equiv \mathbb{R}^2$. Heckman and Vytlacil (2005), Heckman et al. (2006), Vytlacil (2002), and Vytlacil (2006) discuss why the unobserved component in latent index models must be additively separable to achieve identification with an instrumental variable in the presence of essential heterogeneity.

Households face a constrained optimization problem, in that they cannot attend a school below the 0th percentile of quality or live in a neighborhood above the 100th percentile of quality. As a result, actual choices depend on the latent indexes as

$$q_{Si} = \begin{cases} 0 + \varepsilon^0_{Si} & \text{if } q^*_{Si} < 0 \quad \text{where} \quad \varepsilon^0_{Si} \sim \gamma^0_S \text{Beta}(\alpha^0_S, \beta^0_S); \\ q^*_{Si} & \text{if } q^*_{Si} \in (0, 100); \\ 100 - \varepsilon^{100}_{Si} & \text{if } q^*_{Si} > 100 \quad \text{where} \quad \varepsilon^{100}_{Si} \sim \gamma^{100}_S \text{Beta}(\alpha^{100}_S, \beta^{100}_S); \end{cases} \tag{4}$$

and

$$q_{Ni} = \begin{cases} 0 + \varepsilon^0_{Ni} & \text{if } q^*_{Ni} < 0 \quad \text{where} \quad \varepsilon^0_{Ni} \sim \gamma^0_N \text{Beta}(\alpha^0_N, \beta^0_N); \\ q^*_{Ni} & \text{if } q^*_{Ni} \in (0, 100); \\ 100 - \varepsilon^{100}_{Ni} & \text{if } q^*_{Ni} > 100 \quad \text{where} \quad \varepsilon^{100}_{Ni} \sim \gamma^{100}_N \text{Beta}(\alpha^{100}_N, \beta^{100}_N). \end{cases} \tag{5}$$

The ε's can be thought to represent frictions in the housing market.

Discrete levels of school quality

$$D_S = \begin{cases} 1 & \text{if } q_S \in [\underline{q}^D_{S1}, \overline{q}^D_{S1}); \\ \vdots & \quad \vdots \\ J_S & \text{if } q_S \in [\underline{q}^D_{SJ_S}, \overline{q}^D_{SJ_S}) \end{cases} \tag{6}$$

and neighborhood quality

$$D_N = \begin{cases} 1 & \text{if } q_N \in [\underline{q}^D_{N1}, \overline{q}^D_{N1}); \\ \vdots & \quad \vdots \\ J_N & \text{if } q_N \in [\underline{q}^D_{NJ_N}, \overline{q}^D_{NJ_N}) \end{cases} \tag{7}$$

determine potential outcomes as

A2: $Y_{ij_Sj_N} = \mu_{j_Sj_N}(\boldsymbol{X_i}) + U_{ij_Sj_N}$ for $j_S = 1, \ldots, J_S$ and $j_N = 1, \ldots, J_N$.

I add independence, monotonicity and relevance, and integrability assumptions

A3: $(\boldsymbol{X_i}, V_{Si}, V_{Ni}, U_{ij_Sj_N}) \perp\!\!\!\perp \mathbf{Z}_i$ for all $j_S = 1, \ldots, J_S$ and $j_N = 1, \ldots, J_N$

A4: Either

(a) $\boldsymbol{\mu_{S,N}}(\boldsymbol{X_i}, Z_i = 1) \succeq \boldsymbol{\mu_{S,N}}(\boldsymbol{X_i}, Z_i = 0)$ for all i with

$$
\begin{aligned}
(ai) \quad & \mu_S(\boldsymbol{X_i}, Z_i = 1) > \mu_S(\boldsymbol{X_i}, Z_i = 0) \quad \text{for at least one } i \text{ and} \\
& \mu_N(\boldsymbol{X_i}, Z_i = 1) \geq \mu_N(\boldsymbol{X_i}, Z_i = 0) \quad \text{for all } i \text{ or;} \\
(aii) \quad & \mu_S(\boldsymbol{X_i}, Z_i = 1) \geq \mu_S(\boldsymbol{X_i}, Z_i = 0) \quad \text{for all } i \text{ and} \\
& \mu_N(\boldsymbol{X_i}, Z_i = 1) > \mu_N(\boldsymbol{X_i}, Z_i = 0) \quad \text{for at least one } i;
\end{aligned}
$$

or

(b) $\boldsymbol{\mu_{S,N}}(\boldsymbol{X_i}, Z_i = 1) \preceq \boldsymbol{\mu_{S,N}}(\boldsymbol{X_i}, Z_i = 0)$ for all i with

$$
\begin{aligned}
(bi) \quad & \mu_S(\boldsymbol{X_i}, Z_i = 1) < \mu_S(\boldsymbol{X_i}, Z_i = 0) \quad \text{for at least one } i \text{ and} \\
& \mu_N(\boldsymbol{X_i}, Z_i = 1) \leq \mu_N(\boldsymbol{X_i}, Z_i = 0) \quad \text{for all } i \text{ or;} \\
(bii) \quad & \mu_S(\boldsymbol{X_i}, Z_i = 1) \leq \mu_S(\boldsymbol{X_i}, Z_i = 0) \quad \text{for all } i \text{ and} \\
& \mu_N(\boldsymbol{X_i}, Z_i = 1) < \mu_N(\boldsymbol{X_i}, Z_i = 0) \quad \text{for at least one } i;
\end{aligned}
$$

A5: $\mathbb{E}[\,|Y_{j_Sj_N}|\,] < \infty$ for all $j_S = 1, \ldots, J_S$ and $j_N = 1, \ldots, J_N$

The notation $\succeq$ in A4 denotes the two-dimensional partial order where $(q_S^{''}, q_N^{''}) \succeq (q_S^{'}, q_N^{'}) \iff q_S^{''} > q_S^{'}$ and $q_N^{''} > q_N^{'}$. Further research can investigate whether tests of A3 and A4 generalize from those already established in the one-dimensional case (Angrist and Imbens (1995), Huber and Mellace (2015), Mourifié and Wan (2016)). As discussed in Heckman and Vytlacil (2005), A3 and A4 are necessary for identification but not for the definition of causal effects. The assumptions made about potential outcomes through the U_{ij} are that the unobserved term is additive (A2), the instrument is valid (A3), and the potential outcomes have finite mean(A5). No conditional independence assumption is made on the difference between potential outcomes through the joint distribution of unobservables.

The causal parameters of interest here are j to $j+1$ transition-specific Local Average Treatment Effects (LATEs) for the experimental MTO voucher from increasing school quality, neighborhood

quality, or both:

$$\triangle^{LATE}_{j_S+1}(\,Z\,) \equiv \mathbb{E}\big[\,Y_{j_S+1 j_N} - Y_{j_S j_N} \;\big|\; D_S(Z=0)=j_S, D_N(Z=0)=j_N, \; D_S(Z=1)=j_S+1, D_N(Z=1)=j_N\,\big] \tag{8}$$

$$\triangle^{LATE}_{j_N+1}(\,Z\,) \equiv \mathbb{E}\big[\,Y_{j_S j_N+1} - Y_{j_S j_N} \;\big|\; D_S(Z=0)=j_S, D_N(Z=0)=j_N, \; D_S(Z=1)=j_S, D_N(Z=1)=j_N+1\,\big] \tag{9}$$

$$\triangle^{LATE}_{j_S+1 j_N+1}(\,Z\,) \equiv \mathbb{E}\big[\,Y_{j_S+1, j_N+1} - Y_{j_S j_N} \;\big|\; D_S(Z=0)=j_S, D_N(Z=0)=j_N, \; D_S(Z=1)=j_S+1, D_N(Z=1)=j_N+1\,\big] \tag{10}$$

These parameters are instrument-specific, maintaining this distinguishing feature of one-dimensional LATE parameters.

6 Identification of School and Neighborhood Effects

6.1 First Stage: Identifying the School and Neighborhood Choice Model

The goal from the first stage - estimating the school and neighborhood choice model - is to identify those households who would live in different treatment levels depending on whether they receive a voucher or not. Focusing on the school dimension alone for the sake of exposition, this means children that would be in levels (j_S, j_N) without receiving a voucher but in levels (j_S+1, j_N) with a voucher.

The model of school and neighborhood choice can be estimated non-parametrically for unconstrained households. These are households for which $(q^*_{Si}, q^*_{Ni}) \in (0,100) \times (0,100)$. Recalling Equations 4 and 5 along with the fact that the Beta distribution has support [0,1], we can be sure that a household is unconstrained if

$$(q_{Si}, q_{Ni}) \in (\gamma^0_S, 100 - \gamma^{100}_S) \times (\gamma^0_N, 100 - \gamma^{100}_N). \tag{11}$$

For the remainder of the identification analysis suppose that we are focused only on the unconstrained subsample. Then assuming a set of K binary X_i, let $\boldsymbol{X_i}$ be an indicator for being in one of 2^K cells.[7] Alternatively, one might also refer to $\boldsymbol{X_i} = k \in \{1, \ldots, 2^K\}$. After estimating the observed component as

$$\widehat{\boldsymbol{\mu}}_{\boldsymbol{S,N}}(\boldsymbol{X_i}, Z_i) = \mathbb{E}[(q_{Si}, q_{Ni})|(\boldsymbol{X_i}, Z_i)],$$

the unobserved component can be identified as

$$\widehat{\boldsymbol{V_i}} = (q_{Si}, q_{Ni}) - \widehat{\boldsymbol{\mu}}_{\boldsymbol{S,N}}(\boldsymbol{X_i}, Z_i).$$

Finding those induced by the instrument to move would require finding observed characteristics

[7] The generalization to discrete X_i of cardinality greater than 2 is straightforward.

$\boldsymbol{X_i} \in \{1, \ldots, 2^K\}$ and unobserved characteristics $\widehat{\boldsymbol{V_i}}$ such that

$$\widehat{\boldsymbol{\mu}}_{\boldsymbol{S},\boldsymbol{N}}(\boldsymbol{X_i} = k, Z_i = 0) + \widehat{\boldsymbol{V_i}} \in [\underline{q}^D_{Sj}, \overline{q}^D_{Sj}) \times [\underline{q}^D_{Nj}, \overline{q}^D_{Nj}) \text{ and}$$
$$\widehat{\boldsymbol{\mu}}_{\boldsymbol{S},\boldsymbol{N}}(\boldsymbol{X_i} = k, Z_i = 1) + \widehat{\boldsymbol{V_i}} \in [\underline{q}^D_{Sj}, \overline{q}^D_{Sj+1}) \times [\underline{q}^D_{Nj}, \overline{q}^D_{Nj}).$$

This is the identification support set for the LATE in Equation 8, and it is denoted by

$$\mathcal{I}_{j_S+1} = \{\boldsymbol{X_i}, \boldsymbol{V_i} \mid \big(D_S(Z=0) = j_S, \quad D_N(Z=0) = j_N\big),$$
$$\big(D_S(Z=1) = j_S + 1, \ D_N(Z=1) = j_N\big)\}$$

with the notation

$$i \in \mathcal{I}_{j_S+1} \equiv (\boldsymbol{X_i}, \boldsymbol{V_i}) \in \mathcal{I}_{j_S+1}.$$

This set can be found for each j_S to $j_S + 1$ transition by finding the region in $\mathcal{M} \times \mathcal{V}$ where children without vouchers are located in j_S and children with vouchers are located in $j_S + 1$. That is, if

$$\mathcal{I}_{j_S}(Z=0) \equiv \big\{\boldsymbol{X_i}, \boldsymbol{V_i} | \big(D_S(Z=0) = j_S, D_N(Z=0) = j_N\big)\big\}$$

and

$$\mathcal{I}_{j_S+1}(Z=1) \equiv \big\{\boldsymbol{X_i}, \boldsymbol{V_i} | \big(D_S(Z=1) = j_S + 1, \ \ D_N(Z=1) = j_N\big)\big\},$$

then

$$\mathcal{I}_{j_S+1} = \mathcal{I}_{j_S}(Z=0) \cap \mathcal{I}_{j_S+1}(Z=1).$$

6.2 Second Stage: Estimating LATEs

Applying the Wald estimator to children in $\mathcal{I}_{j_S+1}$ identifies the LATE in Equation 8:

$$\triangle^{LATE}_{j_S+1}(\ Z\ ) = \frac{\mathbb{E}[Y | i \in \mathcal{I}_{j_S+1}, Z=1] - \mathbb{E}[Y | i \in \mathcal{I}_{j_S+1}, Z=0]}{\mathbb{E}[D_S | i \in \mathcal{I}_{j_S+1}, Z=1] - \mathbb{E}[D_S | i \in \mathcal{I}_{j_S+1}, Z=0]}.$$

Analogous sets identify the LATEs in Equations 9 and 10. Note the denominator in the Wald estimator for the LATE

$$\triangle^{LATE}_{j_S+1,j_N+1}(\ Z\ ) = \frac{\mathbb{E}[Y | i \in \mathcal{I}_{j_S+1,j_N+1}, Z=1] - \mathbb{E}[Y | i \in \mathcal{I}_{j_S+1,j_N+1}, Z=0]}{Pr[D_S = j_S+1, D_N = j_N+1 | i \in \mathcal{I}_{j_S+1,j_N+1}, Z=1] - Pr[D_S = j_S+1, D_N = j_N+1 | i \in \mathcal{I}_{j_S+1,j_N+1}, Z=0]}.$$

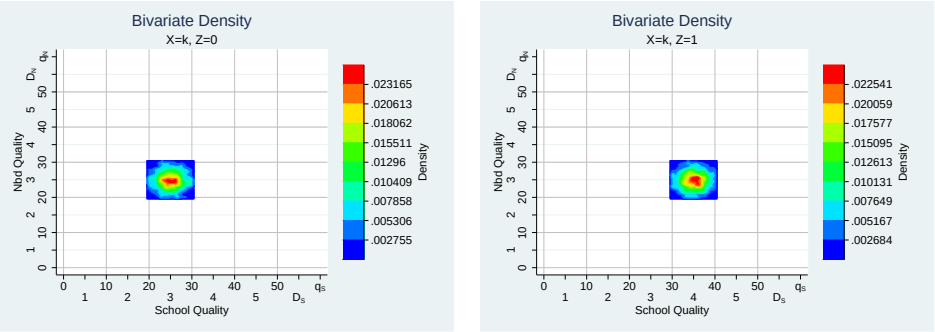
6.3 Discussion

In the absence of an instrument, researchers would typically use a rich set of covariates to ensure that a conditional independence assumption holds. Since we have a randomized instrument and are interested in identifying *Local* Average Treatment Effects, here we instead use covariates for the sake of finding moves across specific margins of treatment. For example, we might find that boys

in Baltimore whose parent had a high school diploma nearly all were in treatment levels (j_S, j_N) without a voucher but nearly all selected into $(j_S + 1, j_N)$ with a voucher. If this were the case, then we could identify a transition-specific LATE for this subgroup.

Figures 6-8 show simulated data illustrating the basic idea of the identification strategy, where both school and neighborhood quality are discretized into deciles for potential outcomes. Suppose that $\boldsymbol{X_i} = k$ denoted girls in Baltimore whose parent had a high school diploma. Figures 6a and 6b show that for this demographic group, no restriction of $\boldsymbol{V_i}$ would be required to identify the LATE of moving from $(j_S, j_N) = (3, 3)$ to $(j_S + 1, j_N) = (4, 3)$ in response to receiving an MTO voucher (Z^M). That is, for Baltimore,

$$\mathcal{I}_{j_S+1=4}(Z^M) = \left\{\boldsymbol{X_i}, \boldsymbol{V_i} | \big(D_S(Z=1) = j_S + 1, \quad D_N(Z=1) = j_N\big)\right\} = \{\boldsymbol{X_i}, \boldsymbol{V_i} | \boldsymbol{X_i} = k\}$$



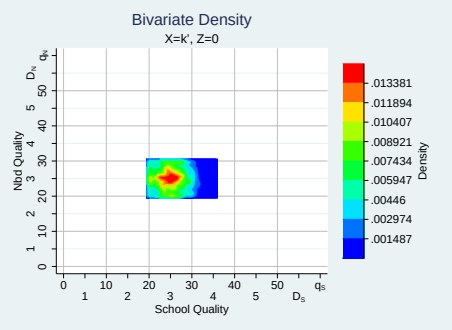
(a) $\boldsymbol{X} = k, Z = 0$ (b) $\boldsymbol{X} = k, Z = 1$

Figure 6: Simulated Data for Demographic Group $\boldsymbol{X} = k$

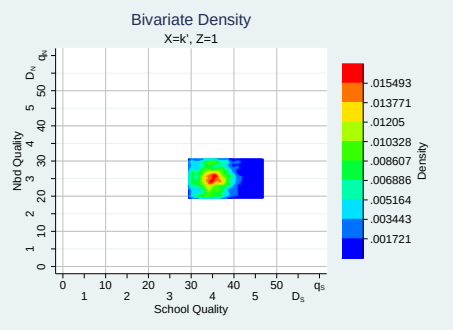
Suppose that $\boldsymbol{X_i} = k'$ denoted boys in Baltimore whose parent had a high school diploma. Figures 7a and 7b show that for this demographic group, some restriction of $\boldsymbol{V_i}$ would be required to identify the LATE of moving from $(j_S, j_N) = (3, 3)$ to $(j_S + 1, j_N) = (4, 3)$. Some share of the lowest V_{Si} (perhaps the lowest 10 percent) would need to be excluded from $\mathcal{I}_{j_S+1=4}$ to identify this LATE in Baltimore:

$$\mathcal{I}_{j_S+1=4}(Z^M) = \left\{\boldsymbol{X_i}, \boldsymbol{V_i} | \big(D_S(Z=1) = j_S{+}1, \quad D_N(Z=1) = j_N\big)\right\} = \{\boldsymbol{X_i}, \boldsymbol{V_i} | \boldsymbol{X_i} = k', V_{Si} \geq Q^{10}(V_{Si})\}$$

We would eliminate V_{Si} below the 10th percentile because the specification of the selection model in Equation A1 implies that households with low values of $\boldsymbol{V_i}$ select into higher quality schools and neighborhoods, and those with high values of $\boldsymbol{V_i}$ select into lower quality schools and neighborhoods.



(a) $\boldsymbol{X} = k', Z = 0$



(b) $\boldsymbol{X} = k', Z = 1$

Figure 7: Simulated Data for Demographic Group $\boldsymbol{X} = k'$

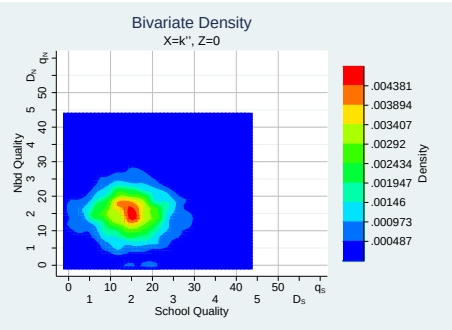
Finally, suppose that $\boldsymbol{X_i} = k''$ denoted black boys in New York City. Figure 8 shows more realistic simulated data that could be used to identify the transition from $(j_S, j_N) = (2, 2)$ to $(j_S+1, j_N+1) = (3, 3)$. If $\boldsymbol{\mu_{S,N}}(\boldsymbol{X_i} = k'', Z_i = 0) = (15, 15)$ and $\boldsymbol{\mu_{S,N}}(\boldsymbol{X_i} = k'', Z_i = 1) = (25, 25)$, then the restrictions on $\boldsymbol{V_i}$ to obtain either

$$\begin{aligned}\mathcal{I}_{j_S=2,j_N=2}(Z^M) &= \left\{\boldsymbol{X_i} = k'', \boldsymbol{V_i} | \left(D_S(Z=0) = 2, \quad D_N(Z=0) = 2\right)\right\} \\ &= \{\boldsymbol{X_i} = k'', \boldsymbol{V_i} | \boldsymbol{\mu_{S,N}}(\boldsymbol{X_i} = k'', Z_i = 0) + \boldsymbol{V_i} \in [10, 20] \times [10, 20]\}\end{aligned}$$

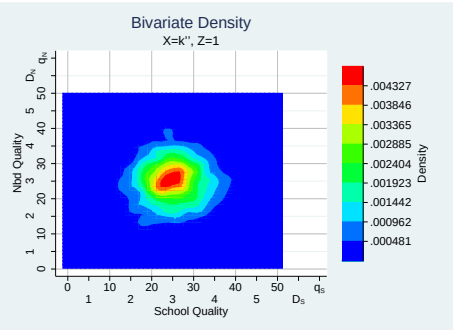
or

$$\begin{aligned}\mathcal{I}_{j_S+1=3,j_N+1=3}(Z^M) &= \left\{\boldsymbol{X_i} = k'', \boldsymbol{V_i} | \left(D_S(Z=1) = 3, \quad D_N(Z=1) = 3\right)\right\} \\ &= \{\boldsymbol{X_i} = k'', \boldsymbol{V_i} | \boldsymbol{\mu_{S,N}}(\boldsymbol{X_i} = k'', Z_i = 1) + \boldsymbol{V_i} \in [20, 30] \times [20, 30]\}\end{aligned}$$

might only require trimming about 20 or 30 percent of the sample with $\boldsymbol{X_i} = k''$.



(a) $\boldsymbol{X} = k'', Z = 0$



(b) $\boldsymbol{X} = k'', Z = 1$

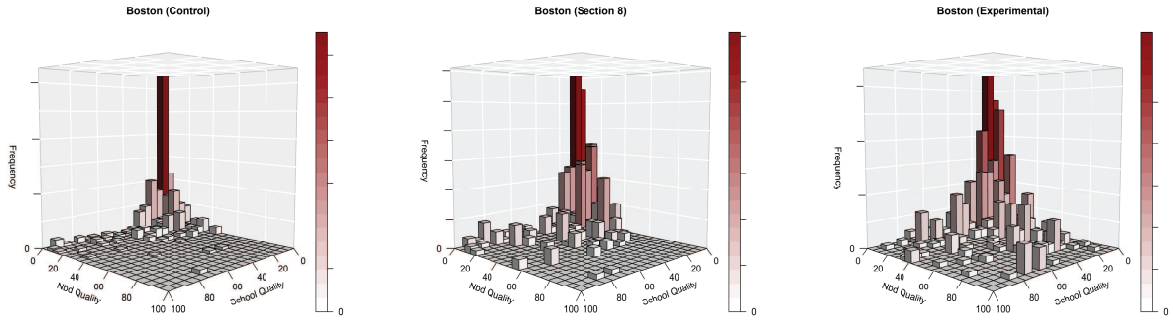
Figure 8: Simulated Data for Demographic Group $\boldsymbol{X} = k''$

7 Empirical Results on School and Neighborhood Effects

7.1 The School and Neighborhood Selection Model

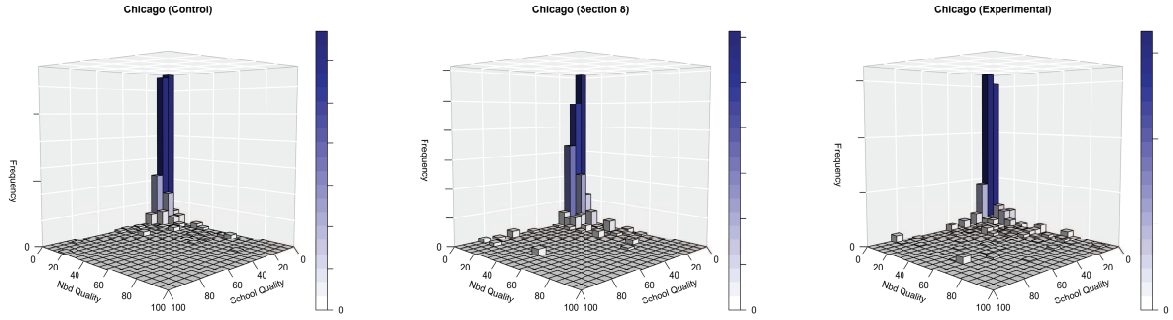
Figures 9-11 show the distribution of children's school and neighborhood quality by site.[8] Some first impressions: Voucher holders in Chicago, whether Section 8 or Experimental, experienced almost no improvement in school or neighborhood quality. Children in Los Angeles were initially attending schools with a range of quality, but almost all living in the worst neighborhoods. The majority of changes in Los Angeles appear to be in neighborhood quality, but there were also improvements in school quality. Vouchers appear to have been most effective at improving school and neighborhood quality across high margins of quality in Boston.

[8] Appendix A displays these data in different formats.



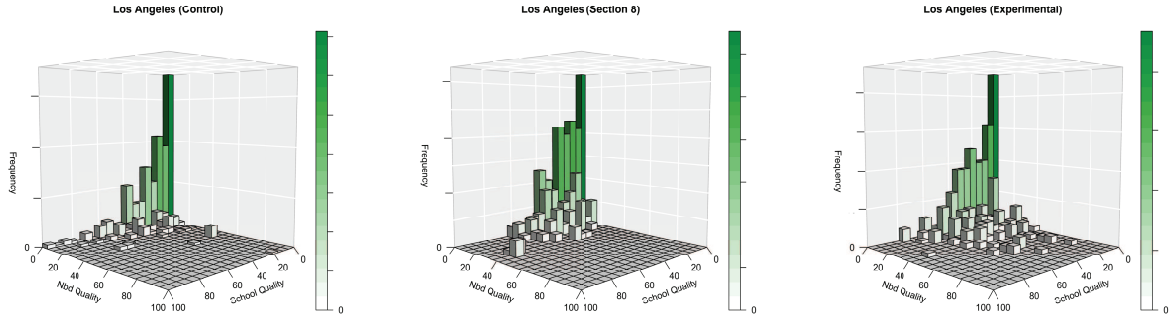
(a) Control (b) Section 8 (c) Experimental

Figure 9: Boston



(a) Control (b) Section 8 (c) Experimental

Figure 10: Chicago



(a) Control (b) Section 8 (c) Experimental

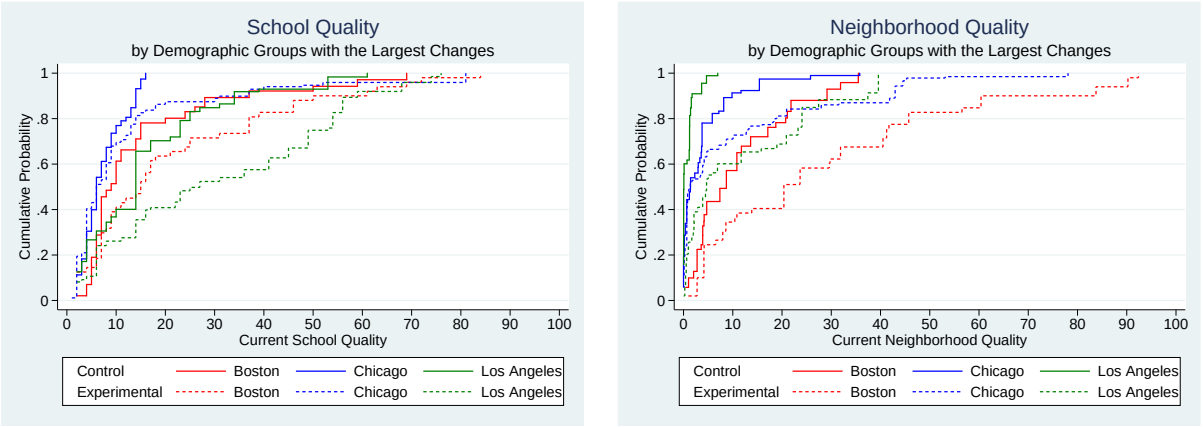
Figure 11: Los Angeles

To find identification support sets $\mathcal{I}$, for each site I do the following: First, calculate the difference

$$\mathbb{E}[q_{Si}|(\boldsymbol{X_i} = k, Z_i = 1)] - \mathbb{E}[q_{Si}|(\boldsymbol{X_i} = k, Z_i = 0)].$$

Second, record the characteristics with the highest and lowest difference in school quality. Last, examine the distributions of school and neighborhood quality for the Control and Experimental voucher groups.

Figures 12a and 12b show that in Boston and Los Angeles there were demographic groups that experienced changes in school quality. Furthermore, these groups experienced changes in neighborhood quality in all three sites.



(a) Largest Changes in School Quality

(b) Largest Changes in Neighborhood Quality

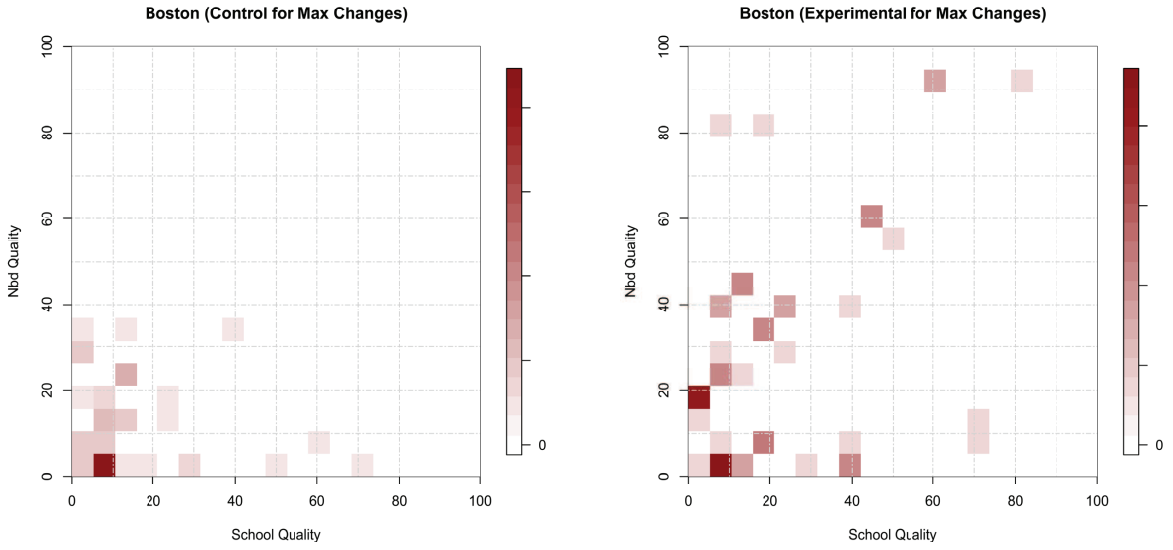
Figure 12: Experimental Voucher and Neighborhood Quality

Unfortunately, sample size and selection patterns are two problems that preclude the estimation of identification support sets, and therefore, the estimation of LATEs. Figure 13 illustrates these problems by showing the joint distribution of school and neighborhood quality for children in Boston. Specifically, the figure shows the Control and MTO Experimental Voucher holders for the demographic subgroup that experienced the largest changes in school quality.

The first problem for identification is illustrated by the right hand side of Figure 13, which shows the weighted frequencies in each 5 percentile by 5 percentile bin. Sample sizes are already small after restricting on the basis of observed characteristics $\boldsymbol{X_i} = k$. Restricting further on the basis of unobserved characteristics $\boldsymbol{V_i}$ would prohibitively decrease sample sizes.

Furthermore, even if sample size were not an issue, most of the control group is near the lower bounds for school and neighborhood quality. For these groups quality will not be in the unconstrained set in Equation 11. Thus it will not be possible to identify the parameters of their choice model.

In other applications it should be possible to use the estimator to identify groups induced by an experiment or instrument to make specific transitions. The best case scenario would be an experiment with many interior individuals and a rich set of observed characteristics. This



(a) Control Group

(b) Experimental Group

Figure 13: Demographic Group with the Largest Changes in School Quality in Boston

estimator could also be used with a future housing mobility program that aimed to vary two or more dimensions of treatment.

8 Conclusion

This paper conducted a subgroup analysis on the effect of receiving a Moving to Opportunity (MTO) housing voucher. The empirical results added some potential sources of voucher effect heterogeneity to those already documented in the literature. To understand the mechanisms driving voucher effect heterogeneity, I specified a joint model of selection into school and neighborhood quality and an estimator to identify these variables' causal effects on math test scores. Unfortunately, I was not able to identify school and neighborhood effects with the MTO data. Nevertheless, I was able to document important correlations between outcomes and changes in school and neighborhood quality.

References

Aliprantis, D. (2012). Redshirting, compulsory schooling laws, and educational attainment. *Journal of Educational and Behavioral Statistics 37*(2), 316–338.

Aliprantis, D. (2016). Human capital in the inner city. *Empirical Economics*. doi: 10.1007/s00181-016-1160-y.

Aliprantis, D. (2017). Assessing the evidence on neighborhood effects from Moving to Opportunity. *Empirical Economics 52*(3), 925–954.

Aliprantis, D. and F. G.-C. Richter (2016). Evidence of neighborhood effects from Moving to Opportunity: LATEs of neighborhood quality. *FRB of Cleveland WP 12-08r*.

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

Barua, R. and K. Lang (2016). School entry, educational attainment, and quarter of birth: A cautionary tale of a Local Average Treatment Effect. *Journal of Human Capital 10*(3), 347–376.

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., J. Hyman, and M. Kapustin (2017). Predictors of successful housing voucher lease-up and implications for estimated labor market responses. *Mimeo., University of Connecticut*.

Davis, M., J. Gregory, D. Hartley, and K. Tan (2017). Neighborhood choices, neighborhood effects and housing vouchers. *FRB of Chicago Working Paper WP-2017-2*.

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.

Fruehwirth, J. C. (2014). Can achievement peer effect estimates inform policy? A view from inside the black box. *The Review of Economics and Statistics 96*(3), 514–523.

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

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

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

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

Huber, M. and G. Mellace (2015). Testing instrument validity for LATE identification based on inequality moment constraints. *The Review of Economics and Statistics 97*(2), 398–411.

Imai, K. and M. Ratkovic (2013, 03). Estimating treatment effect heterogeneity in randomized program evaluation. *Ann. Appl. Stat. 7*(1), 443–470.

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

Kling, J. R., J. B. Liebman, and L. F. Katz (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.

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.

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.

Mourifié, I. and Y. Wan (2016). Testing Local Average Treatment Effect assumptions. *The Review of Economics and Statistics*. doi:10.1162/REST_a_00622.

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. (2015). 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.

Rockoff, J., R. Chetty, and J. Friedman (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review 104*(9), 2593–2632.

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.

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.

Tian, L., A. A. Alizadeh, A. J. Gentles, and R. Tibshirani (2014). A simple method for estimating interactions between a treatment and a large number of covariates. *Journal of the American Statistical Association 109*(508), 1517–1532.

Tincani, M. (2015). Heterogeneous peer effects and rank concerns: Theory and evidence. *Mimeo., University College London*.

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.

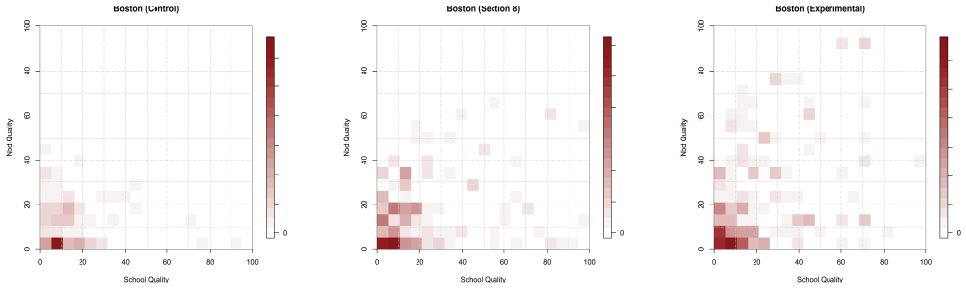
Vytlacil, E. (2006). Ordered discrete-choice selection models and local average treatment effect assumptions: Equivalence, nonequivalence, and representation results. *The Review of Economics and Statistics 88*(3), 578–581.

Wager, S. and S. Athey (2017). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*. doi: 10.1080/01621459.2017.1319839.

Zuberi, A. (2010). Limited exposure: Children's activities and neighborhood effects in the Gautreaux Two housing mobility program. *Journal of Urban Affairs 32*(4), 405–423.

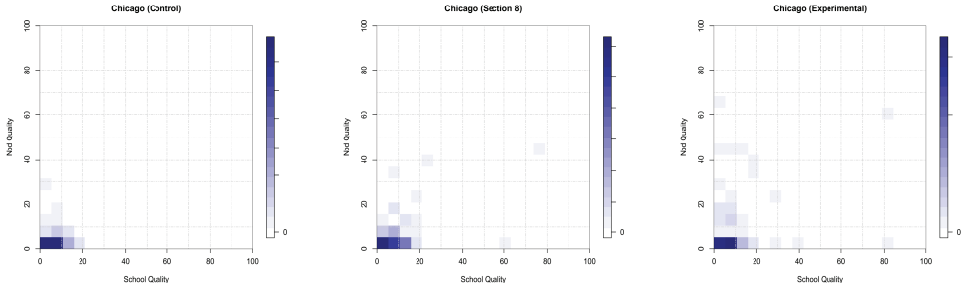
A Additional Figures

Selection into school and neighborhood quality by site and voucher type is shown in the figures below.



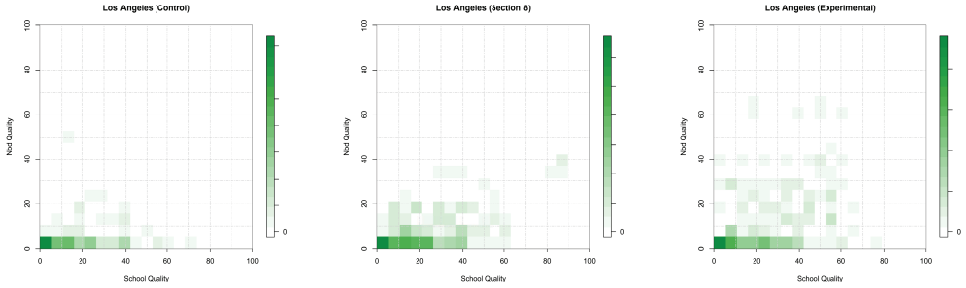
(a) Control (b) Section 8 (c) Experimental

Figure 14: Boston



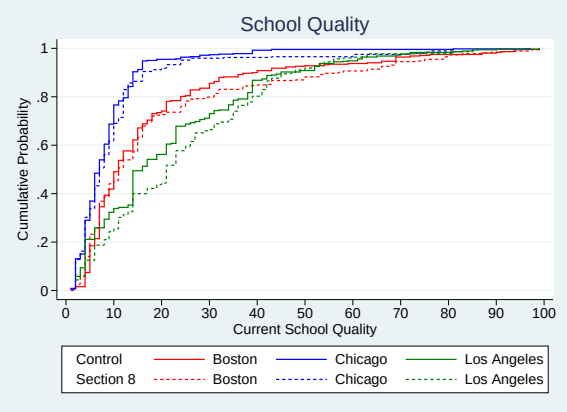
(a) Control (b) Section 8 (c) Experimental

Figure 15: Chicago

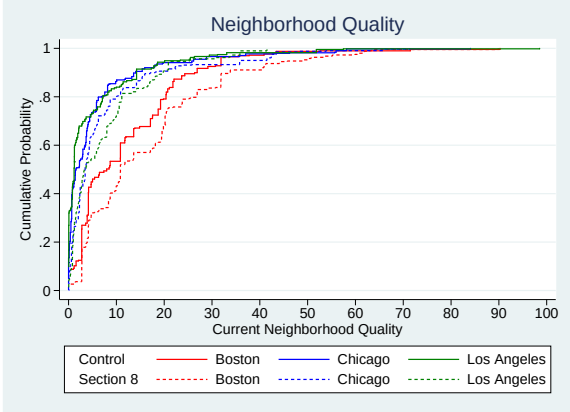


(a) Control (b) Section 8 (c) Experimental

Figure 16: Los Angeles

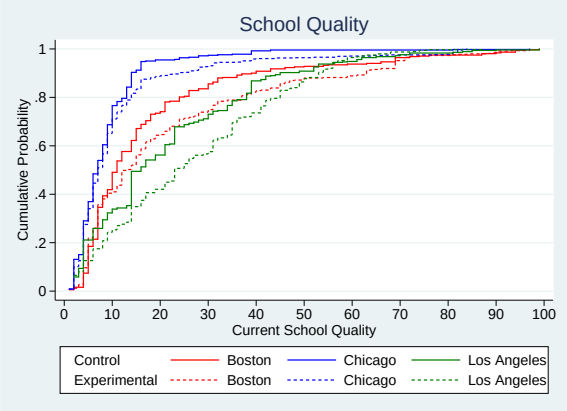


(a) School Quality

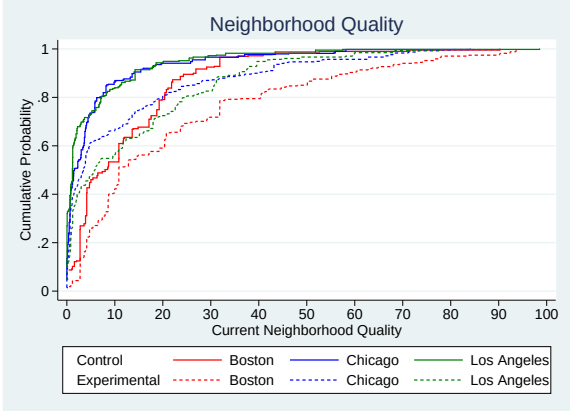


(b) Neighborhood Quality

Figure 17: Section 8



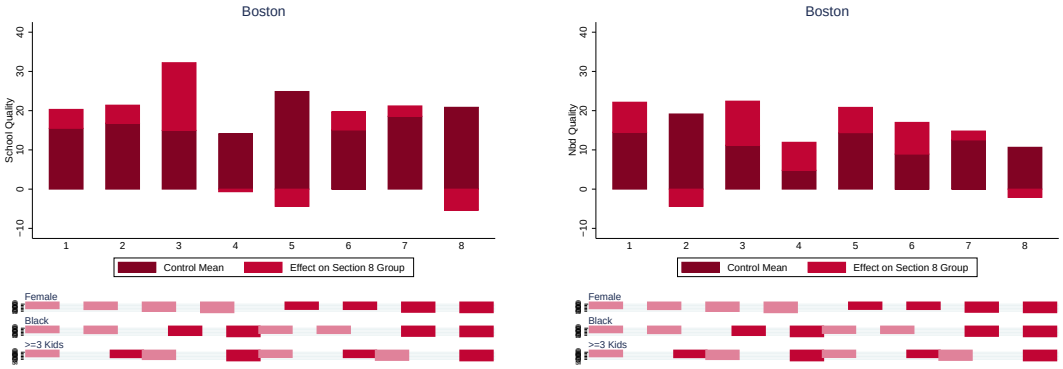
(a) School Quality



(b) Neighborhood Quality

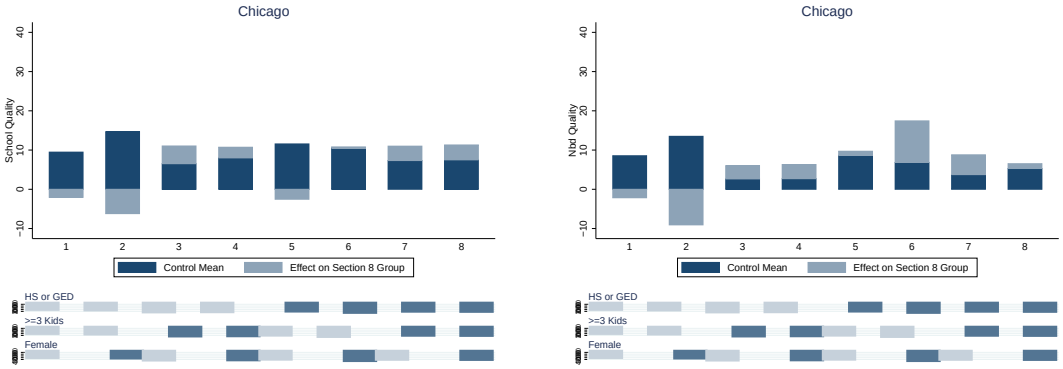
Figure 18: Experimental

B Section 8 Analysis



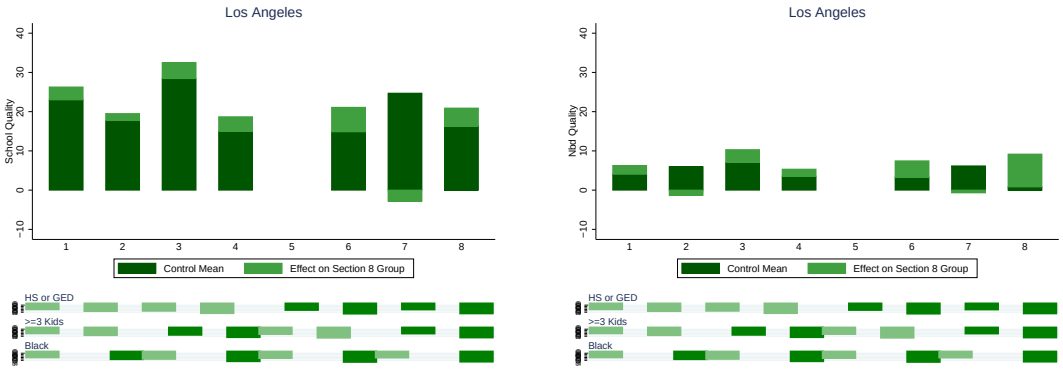
(a) School Quality (b) Neighborhood Quality

Figure 19: Boston by Demographic Group



(a) School Quality (b) Neighborhood Quality

Figure 20: Chicago by Demographic Group



(a) School Quality (b) Neighborhood Quality

Figure 21: Los Angeles by Demographic Group

C Unconstrained Regressions

Table 2: Effect Heterogeneity

Observed Characteristic	σ	t-statistic	P-Value
$\geq$3 Kids in HH in Boston			
Coefficient on Z^M	0.03	0.30	0.763
	(0.09)		
Coefficient on Interaction Term	–0.21	–1.90	0.058
	(0.11)		
Female in Chicago			
Coefficient on Z^M	–0.00	–0.03	0.972
	(0.09)		
Coefficient on Interaction Term	0.27	2.71	0.007
	(0.10)		
Black in Los Angeles			
Coefficient on Z^M	0.18	1.83	0.068
	(0.10)		
Coefficient on Interaction Term	–0.53	–5.01	0.000
	(0.11)		

Note: Site-specific least-squares regressions of

$$T_i = \beta_0 + \beta_3 Z_i^M + \beta_4 Z_i^M \times \mathbf{1}\{\boldsymbol{X_i} = k\} + \epsilon_{Ti}$$

where T_i is WJR Math Calculation subtest score, $Z_i^M = 1$ for MTO voucher holders and $Z_i^M = 0$ for the control group, and $\boldsymbol{X_i} = k$ is the observed characteristic specified in the table. All regressions are weighted, P-Value is from a two-tailed test.