Chapter 7
Neighborhood Effects

When characterizing the location choice problem of households, we have emphasized that residential amenities play an important role in the problem. Our economic geography and system of cities frameworks have shown that people have preferences over these amenities, and sort according to them.

We have also argued that there may be residential agglomeration externalities. People may like to be near other people, and there are many reasons why this may be the case. You may like to have a high skill, high income neighbor, in the hope that you too, will become high skilled and rich. At the same time, people may not want to be near other people if the agglomeration externalities are negative, in the form of congestion externalities. More people around you means more traffic, more noise, more diseases, etc ...

The estimates from Ahlfeldt et al. (2015) suggest that the positive benefits of residential agglomeration outset the costs. We would like to know why this is the case. Why are places and neighbors so important in determining the utility of workers, such that they determine location choices?

We will start this chapter outlining some theoretical reasons why these neighborhood effects may exist. Then, we will turn to empirical evidence on this neighborhood effects. This is a difficult empirical problem to tackle with observational data, and we will see why. So we will look at evidence from a large scale experiment: Moving to Opportunity (MTO). Roughly speaking, this program relocated people from "bad" neighborhoods to 'good' neighborhoods. We will see how to estimate neighborhood effects in this experimental framework, and what the effects were.

This chapter follows Lawrence Katz's lectures on labor economics quite closely, with departures in particular sections.

7.1 Do Neighborhoods Matter?

There are correlations between neighborhoods and individual outcomes. Children who grow up in bad neighbourhoods tend to do worse later in life. This can not be

taken as evidence of a causal effect of neighborhoods, because of sorting. Maybe the families where children would do badly later in life, regardless of where they locate, decided to locate in the bad neighborhoods.

There are some models about how neighbors may affect individual behavior (Mayer and Jencks, 1989):

- Disadvantaged Neighbors are a Disadvantage (Contagion): Individual outcomes correlate positively with neighbours' outcomes. So children surrounded by children with high propensity to crime will tend to commit crimes themselves. Children surrounded by high-achieving children may also become high achieving.
- Advantaged Neighbors Are A Disadvantage (Relative Deprivation): Individual outcomes correlate negatively with neighbours' outcomes. For example, a child may become frustrated from having high-achieving peers and exert little effort in school.
- Disadvantaged Neighbors Are Irrelevant: Individual outcomes do not correlate with neighbor's outcomes.
- Neighbors don't Matter but Neighborhoods Do: Even if neighbors per se do not directly affect behavior, neighborhood institutions, resources, and physical environment do.

7.1.1 Contagion effects

Under the hypothesis of contagion effects, the outcomes of an individual are positively correlated with the outcomes of their neighbors. Measuring the extent of these contagion effects is important, but quite difficult at the same time.

Some mechanisms behind contagion effects are the following:

- Physical Externalities: If more people are involved in a crime, it is tougher for the police to catch any individual.
- Network Externalities: Having a network is valuable, and the value of the network is proportional to its size.
- Social Learning / Information: Neighbors and peers provide information on how to sign up for welfare, avoid police, adopt a new technology.
- Pure Preference Externalities: Individuals get direct pleasure from being "in" and doing what your peers do.
- Stigma Effects: Negative signal from delinquent behavior declines when more peers and neighbors are doing it.

Contagion effects are important because they change the multipliers of social policy. Suppose that there is a place-based policy B_r that increases the welfare of individual i in neighborhood r, W_{ir}. B_{ir} measures how much individual i benefits from the policy. Simultaneously, suppose that the welfare of an individual depends on the average welfare in the neighborhood, W_r.

$$W_{ir} = B_{ir}\alpha + W_r\beta \tag{7.1}$$

Averaging and solving for W_r, we see that the overall group mean welfare depends on the group mean benefit level:

$$W_r = B_r\alpha/(1-\beta) \tag{7.2}$$

The effect of increasing the benefit level is mediated by the the neighborhood effect:

$$dW_r/dB_r = \alpha/(1-\beta) \tag{7.3}$$

Because of the social multiplier arising from peer effects, changes in average welfare benefits have a larger effect then one would infer by looking at individual at the effects of individual level variation in benefits on welfare within the reference group.

The reason why these contagion effects are hard to measure is because correlation between individual and neighbor outcomes is not quite informative about them. There are many other reasons why outcomes may be correlated within neighborhoods. For example, individual behavior may not be affected by the neighbors' outcomes, but by their characteristics. For example, a kid may learn more near other kids who are studying, but maybe he would learn from them regardless of whether they did well on the test. Effects that come from the neighbor's characteristics, and not from the neighbors' outcomes, are called contextual effects.

Another reasons why there may be correlation within neighborhoods may be:

- Neighborhood characteristics (correlated effects) directly affect outcomes,
- Correlated unobservable individual attributes, like unobserved individual or family factors (parental income, attitudes and skills) may be correlated within a neighborhood.

One could write a general model that allows for both contagion and contextual effects. Let y_{in} be the outcome for an individual in neighborhood n. y_n is the mean outcome in neighborhood n. z are characteristics of the neighbors, and x are characteristics of the neighborhood. The model is:

$$y_{in} = \alpha + z_{in}\gamma + y_n\beta + z_n\pi + x_n\delta + \varepsilon_{in} \tag{7.4}$$

Here, γ measures contextual effects, β measures contagion effects, and δ measures neighborhood effects. Without exogenous variation in each of these three sources, it is hard to separate them.

7.2 Moving to Opportunity

The Moving to Opportunity Experiment (MTO), examined in Kling et al. (2007) and many other papers, tackled estimation of neighborhood effects using the gold

standard: experimental variation. The program targeted families with children living in public housing in high poverty neighborhoods, defined as census tracts with a poverty rate larger than 40%. It ran from 1994 to 1997 in six cities: Baltimore, Boston, Chicago, Los Angeles and New York.

Participants were offered vouchers to relocate to other neighborhoods. They were randomly assigned to three groups:

- Control group: No voucher was offered, but the households remain eligible for public housing.
- Section 8: Offered Conventional Section 8 Vouchers, that paid any rent over 30% of the household's income, up to 15000 dollars.
- Experimental: Offered a Restricted Section 8 Voucher, that could only be used to relocate to a low-poverty area ($< 10\%$ poverty rate). It also included mobility counseling.

We will review how to get causal effects in this randomized experiment framework.

7.2.1 The Program Evaluation Problem

We will outline the problem with a potential outcomes notation. Angrist and Pischke (2008) provide a detailed version of this.

An individual i can be in either a treated state "1" or an untreated state "0". Denote:

- Y_{0i} is the outcome for i without the treatment or program
- Y_{1i} is the outcome for i with the treatment or program
- $d_i = 1$ if i receives the treatment and 0 if it does not

The fundamental problem is that we do not observe both Y_{0i} and Y_{1i}. We only observe the treatment outcome for treatment individuals, and the control outcome for control individuals. We can write the observed outcome as:

$$Y_i = d_i Y_{1i} + (1 - d_i) Y_{0i} \tag{7.5}$$

The causal effect from treatment for an individual i is :

$$\alpha_i = Y_{1i} - Y_{0i} \tag{7.6}$$

Which is unobservable, because we do not observe both Y_{0i} and Y_{1i}. Although this in unobservable, we may not be interested in this. Instead of the individual effect, we would like to know what was the average effect of the program across individuals.

The Average Treatment Effect (ATE) is the expected effect for a randomly selected person from entire population.

$$ATE = E[Y_{1i} - Y_{0i}] \tag{7.7}$$

The Treatment on the Treated (TOT), also known as the Average Treatment Effect on the Treated, or the Selected Average Treatment Effect , is the average effect among the treated group:

$$TOT = E[Y_{1i} - Y_{0i}|di = 1] = E[\alpha_i|d_i = 1] \tag{7.8}$$

The standard approach is to compare the mean post-program outcome of the treatment group and the comparison group :

$$\begin{aligned}
T\hat{O}T &= E[Y_i|d_i = 1] - E[Y_i|d_i = 0] \tag{7.9}\\
&= E[Y_{1i}|d_i = 1] - E[Y_{0i}|d_i = 1] + (E[Y_{0i}|d_i = 1] - E[Y_{0i}|d_i = 0])\\
&= E[\alpha_i|d_i = 1] + (E[Y_{0i}|d_i = 1] - E[Y_{0i}|d_i = 0])
\end{aligned}$$

This decomposition highlights the selection bias problem in estimating TOT. The first term is the parameter of interest, but the term in brackets is the selection bias term. If assignment to the treatment is nonrandom, then omitted variables that affect both Y_{0i} and selection into the program will generate selection bias. Selection bias arises when the non-participants differ from the participants in the non-participant state.

7.2.2 Randomized Social Experiment with Full Compliance

Randomization of the treatment solves the selection bias problem. By randomizing, this approach generates a control group consisting of those persons who would have participated but were randomly denied treatment.

Under the assumption of no randomization bias –so that randomization per se does not change the pool of applicants or the operation of the program – the treatment group mean outcome provides an estimate of $E[Y_1|d = 1]$ and the control group mean outcome provides an estimate of $E[Y_0|d = 1]$.

Because of random assignment $E[Y_{0i}|d_i = 1] = E[Y_{0i}|d_i = 0] = E[Y_0i]$. So the selection bias term cancels out and the difference in means provides an unbiased estimate of TOT.

The experiment supplements missing data by providing an estimate of $E[Y_0|d = 1]$ from the sample mean for the control group and $E[Y_1|d = 1]$ from the mean of the treatment group (which is also available in observational studies). The $TOT = E(\alpha|d = 1)$ and the distributions of outcomes with and without treatment can be estimated for the $d = 1$ group.

But one can't recover the overall distribution of treatment effects $F(\alpha)$ without strong additional assumptions.

7.2.3 Eligibility Randomization Experiments and Partial Compliance

Eligibility for a program (a treatment) is randomized (typically among applicants to the program) and then the eligibles "choose" whether to participate; those randomized out of eligibility (the control group) can't participate. If some eligibles don't get treatment (don't comply), then one has a problem of partial (or incomplete) compliance. Eligibility randomization is the usual approach for most real world clinical drug trials, social experiments, and field experiment.

Eligibility randomization allows one to directly estimate the mean effect of eligibility for the program on the population included in the experiment.

Consider a population of persons normally eligible for a program. Let $Z = 1$ if the person is kept eligible after randomization and $Z = 0$ if the person loses eligibility.

Eligibility Z is randomly assigned. Actual participation C only equals 1 if the individual is eligible and chooses to participate to participate.

The Intent-to-Treat Effect (ITT) $= E[Y|Z = 1] - E[Y|Z = 0] =$ is the difference in mean outcomes for eligibles and ineligibles if eligibility is randomly assigned.

The TOT can be estimated from an eligibility randomization experiment under the assumptions that:

1. treatment group (eligibility) assignment is truly random;
2. the effect of treatment group assignment on outcomes operates only through participating in the program itself (using a housing voucher in MTO) with no direct effect of eligibility per se;
3. control group members (the ineligibles) cannot participate in the program.

Under these assumptions, the difference in average outcomes of eligibles and ineligibles divided by fraction of eligibles who participate provides an unbiased estimate of the TOT:

$$
\begin{aligned}
TOT &= E[Y_1 - Y_0 | d = 1] \\
&= (E[Y|Z = 1] - E[Y|Z = 0]) / P(C = 1 | Z = 1) \\
&= ITT / P(C = 1 | Z = 1)
\end{aligned} \tag{7.10}
$$

where $P()$ is the probability function, so that $P(C = 1 | Z = 1)$ is the program participation rate.

In the MTO framework, TOT is estimated difference in outcomes between those who actually use the program (MTO voucher) those in the Control group who would have used the program (MTO voucher) if it had been offered to them.

To assess the magnitude of the TOT effect in relative as well as absolute terms, it is useful to have a benchmark level of the outcome in the absence of treatment for comparison. We use the mean outcome for treated compliers and the TOT difference to impute the Control Complier Mean outcome (CCM).

$$CCM = E[Y|C=1, Z=0] \tag{7.11}$$
$$= E[Y|C=1, Z=1] - E[Y|C=1, Z=1] - E[Y|C=1, Z=0]$$
$$= E[Y|C=1, Z=1] - TOT$$

Although $E[Y|C=1, Z=0]$ is not directly observable, $E[Y|C=1, Z=1]$ and TOT can be estimated.

7.2.4 Short-run results of MTO

Kling et al. (2007) provides results of an intermediate evaluation of MTO. Figure 7.1 shows that the program had its intended purpose. Households who received the experimental vouchers moved to neighborhoods that had much lower poverty rates than those in the traditional vouchers group and in the control group.

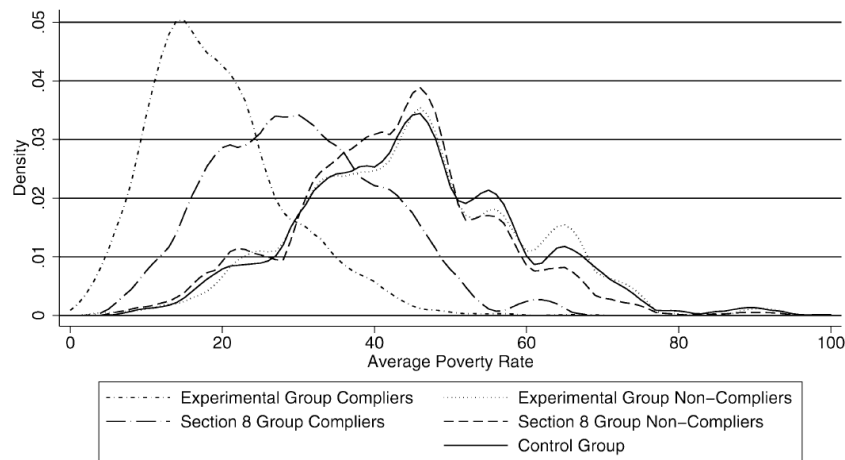**Fig. 7.1** Distribiution of poverty for treatment and control groups.

FIGURE 1.—Densities of average poverty rate, by group. Average poverty rate is a duration-weighted average of tract locations from random assignment through 12/31/2001. Poverty rate is based on linear interpolation of 1990 and 2000 Censuses. Density estimates used an Epanechnikov kernel with a half-width of 2.
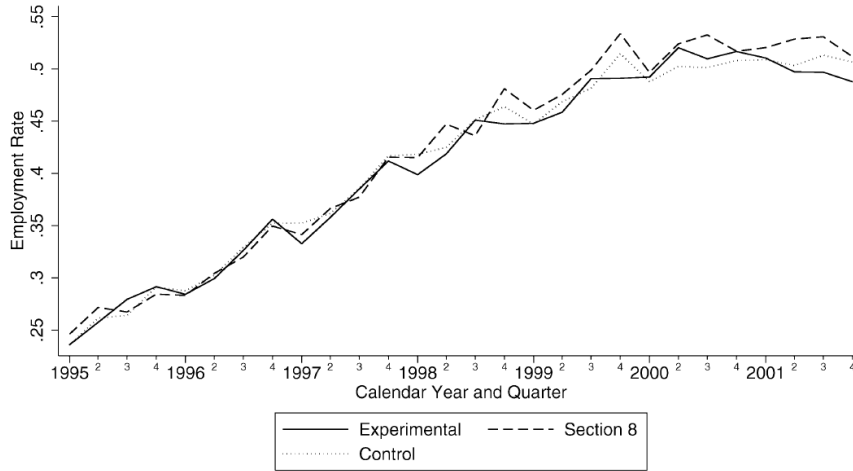
Source: Kling et al. (2007)

Surprisingly, when Kling et al. (2007) looked at the outcomes from the program, they did not find many effects. Table 7.1 shows the subset of variables for which they found estimated effects significant at the 5% level. It is quite surprising how "empty" this table looks. There are not any effects in any economic indicator, such as employment or wages. Impacts tend to be related to stress and mental health, which

they later validate showing a negative correlation between neighborhood poverty rates and mental health indicators.

Figure 7.2 emphasizes the lack of differences in economic status between the treatment and control groups in MTO. The trajectories of employment look quite similar across the groups.

Fig. 7.2 Trajectories of employment for MTO participants by group

Source: Kling et al. (2007)

For a while, the MTO results puzzled many economists. With a crisp-clear identification, there did not seem to be many effects of a neighborhood. Over time, we have learned many reasons why these interim evaluations did not find any effects.

7.2.5 Long run effects of MTO

One potential reason why no effects were found in the original MTO study may have been the difference between short-run and long-run effects. Chetty et al. (2016) follow up on the MTO participants, using confidential tax data. They stress that there are differences between the effects for adults and kids, and that the time of exposure to the better neighborhood may be important in determining the size of the effects. Turns out that all of these differences are essential to understand the extent of neighborhood effects.

First, Chetty et al. (2016) follow the outcomes of individuals who were exposed to MTO when they were children. They separate young children (< 13 when they moved) from teenagers (> 13 when they moved). Table 7.2 shows that all children

Table 7.1 Short-run effects of MTO: Selected variables

SPECIFIC OUTCOMES WITH EFFECTS SIGNIFICANT AT 5 PERCENT LEVEL[a]

	E/S (i)	CM (ii)	ITT (iii)	TOT (iv)	CCM (v)
A. Adult outcomes					
Obese, BMI $\geq$ 30	E − C	0.468	−0.048 (0.022)	−0.103 (0.047)	0.502
Calm and peaceful	E − C	0.466	0.061 (0.022)	0.131 (0.047)	0.443
Psychological distress, K6 z-score	E − C	0.050	−0.092 (0.046)	−0.196 (0.099)	0.150
B. Youth (female and male) outcomes					
Ever had generalized anxiety symptoms	E − C	0.089	−0.044 (0.019)	−0.099 (0.042)	0.164
	S − C	0.089	−0.063 (0.019)	−0.114 (0.035)	0.147
Ever had depression symptoms	S − C	0.121	−0.039 (0.019)	−0.069 (0.035)	0.134
C. Female youth outcomes					
Psychological distress, K6 scale z-score	E − C	0.268	−0.289 (0.094)	−0.586 (0.197)	0.634
Ever had generalized anxiety symptoms	E − C	0.121	−0.069 (0.027)	−0.138 (0.055)	0.207
	S − C	0.121	−0.075 (0.029)	−0.131 (0.051)	0.168
Used marijuana in the past 30 days	E − C	0.131	−0.065 (0.029)	−0.130 (0.059)	0.202
	S − C	0.131	−0.072 (0.032)	−0.124 (0.056)	0.209
Used alcohol in past 30 days	S − C	0.206	−0.091 (0.038)	−0.155 (0.056)	0.306
D. Male youth outcomes					
Serious nonsports accident or injury in past year	E − C	0.062	0.087 (0.026)	0.215 (0.064)	0
	S − C	0.062	0.080 (0.028)	0.157 (0.058)	0
Ever had generalized anxiety symptoms	S − C	0.055	−0.049 (0.024)	−0.098 (0.047)	0.126
Smoked in past 30 days	E − C	0.125	0.103 (0.032)	0.257 (0.084)	0
	S − C	0.125	0.151 (0.037)	0.293 (0.073)	0.014

[a]E/S: indicates whether the row is experimental − control (E − C) or Section 8 − control (S − C). CM, control mean; ITT, intent-to-treat, from Equation (1); TOT, treatment-on-treated, from Equation (2); CCM, control complier mean. Robust standard errors adjusted for household clustering are in parentheses. The estimated equations all include site indicators and the baseline covariates listed in Appendix A with those in Table A1 included for adults and those in Tables A1 and A2 for youth. Rows shown in the table to illustrate magnitudes were selected based on ITT p-values < 0.05 and are 17 of 120 from the set of specific contrasts (E − C, S − C), based on the outcomes (15 for adults and 15 for youth) and subgroups—adults, youth (female and male), female youth, and male youth—described in the notes to Table II.

Source: Kling et al. (2007)

moved to neighborhoods with lower poverty rates. Any difference among these children should not be attributed to ending up in different neighborhoods.

Table 7.2 Decrease in poverty rates for MTO children recipients, separated by age

	Housing voucher take-up (1)	Poverty rate in tract one year post- RA		Mean poverty rate in tract post-RA to age 18		Mean poverty rate in zip post-RA to age 18	
		ITT (2)	TOT (3)	ITT (4)	TOT (5)	ITT (6)	TOT (7)
Panel A. Children < age 13 at random assignment							
Exp. versus control	47.66***	−17.05***	−35.96***	−10.27***	−21.56***	−5.84***	−12.23***
	(1.653)	(0.853)	(1.392)	(0.650)	(1.118)	(0.425)	(0.752)
Sec. 8 versus control	65.80***	−14.88***	−22.57***	−7.97***	−12.06***	−3.43***	−5.17***
	(1.934)	(0.802)	(1.024)	(0.615)	(0.872)	(0.423)	(0.622)
Observations	5,044	4,958	4,958	5,035	5,035	5,035	5,035
Control group mean	0	50.23	50.23	41.17	41.17	31.81	31.81
Panel B. Children age 13–18 at random assignment							
Exp. versus control	40.15***	−14.00***	−34.70***	−10.04***	−24.66***	−5.51***	−13.52***
	(2.157)	(1.136)	(2.231)	(0.948)	(1.967)	(0.541)	(1.113)
Sec. 8 versus control	55.04***	−12.21***	−22.03***	−8.60***	−15.40***	−3.95***	−7.07***
	(2.537)	(1.078)	(1.738)	(0.920)	(1.530)	(0.528)	(0.921)
Observations	2,358	2,302	2,302	2,293	2,293	2,292	2,292
Control group mean	0	49.14	49.14	47.90	47.90	35.17	35.17

Notes: Columns 1, 2, 4, and 6 report ITT estimates from OLS regressions (weighted to adjust for differences in sampling probabilities across sites and over time) of an outcome on indicators for being assigned to the experimental voucher group and the Section 8 voucher group as well as randomization site indicators. Columns 3, 5, and 7 report TOT estimates using a 2SLS specification, instrumenting for voucher take-up with the experimental and Section 8 assignment indicators. Standard errors, reported in parentheses, are clustered by family. Panel A restricts the sample to children below age 13 at random assignment; panel B includes children between age 13 and 18 at random assignment The estimates in panels A and B are obtained from separate regressions. The dependent variable in column 1 is an indicator for the family taking up an MTO voucher and moving. The dependent variable in columns 2 and 3 is the census tract-level poverty rate one year after random assignment. The dependent variable in columns 4–7 is the duration-weighted mean poverty rate in the census tracts (columns 4 and 5) and zip codes (columns 6 and 7) where the child lived from random assignment till age 18. The sample in this table includes all children born before 1991 in the MTO data for whom an SSN was collected prior to RA because we were unable to link the MTO tract-level location information to the tax data. This sample is nearly identical our linked analysis sample because 99.1 percent of the children with nonmissing SSNs are matched to the tax data. The duration-weighted poverty rate is constructed using information on the addresses where the youth lived from random assignment up to their 18th birthday, weighted by the amount of time spent at each address. Census tract poverty rates in each year are interpolated using data from the 1990 and 2000 decennial censuses as well as the 2005–2009 American Community Survey, as in Sanbonmatsu et al. (2011); zip code poverty rates are from census 2000 only and are not interpolated.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Source: Chetty et al. (2016)

Table 7.3 shows that in the long run, MTO had a discernible effect! Young treated children had larger incomes as adults. Over 2008-2012, they earned about 3500 dollars more than the control group. This is a substantial effect. On the other hand, older children do not show any effect, or, if there is any effect, it was negative.

In fact, it seems that MTO may have hurt the older children, suggesting that they had more trouble coping with the change on neighborhoods, and lending some credence to the relative deprivation model of neighborhood effects. Table 7.4 shows

Table 7.3 Treatment effects for economic outcomes among MTO children recipients, separated by age

	W-2 earnings ($) 2008–2012	Individual earnings 2008–2012 ($)			Individual earnings ($)		Employed (%)	Hhold. inc. ($)	Inc. growth ($)
	ITT (1)	ITT (2)	ITT w/ controls (3)	TOT (4)	Age 26 ITT (5)	2012 ITT (6)	2008–2012 ITT (7)	2008–2012 ITT (8)	2008–2012 ITT (9)
Panel A. Children < age 13 at random assignment									
Exp. versus control	1,339.8** (671.3)	1,624.0** (662.4)	1,298.9** (636.9)	3,476.8** (1,418.2)	1,751.4* (917.4)	1,443.8** (665.8)	1.824 (2.083)	2,231.1*** (771.3)	1,309.4** (518.5)
Sec. 8 versus control	687.4 (698.7)	1,109.3 (676.1)	908.6 (655.8)	1,723.2 (1051.5)	551.5 (888.1)	1,157.7* (690.1)	1.352 (2.294)	1,452.4** (735.5)	800.2 (517.0)
Observations	8,420	8,420	8,420	8,420	1,625	2,922	8,420	8,420	8,420
Control group mean	9,548.6	11,270.3	11,270.3	11,270.3	11,398.3	11,302.9	61.8	12,702.4	4,002.2
Panel B. Children age 13–18 at random assignment									
Exp. versus control	−761.2 (870.6)	−966.9 (854.3)	−879.5 (817.3)	−2,426.7 (2,154.4)	−539.0 (795.4)	−969.2 (1,122.2)	−2.173 (2.140)	−1,519.8 (11,02.2)	−693.6 (571.6)
Sec. 8 versus control	−1,048.9 (932.5)	−1,132.8 (922.3)	−1,136.9 (866.6)	−2,051.1 (1,673.7)	−15.11 (845.9)	−869.0 (1213.3)	−1.329 (2.275)	−936.7 (11,85.9)	−885.3 (625.2)
Observations	11,623	11,623	11,623	11,623	2,331	2,331	11,623	11,623	11,623
Control group mean	13,897.1	15,881.5	15,881.5	15,881.5	13,968.9	16,602.0	63.6	19,169.1	4,128.1

Notes: Columns 1–3 and 5–9 report ITT estimates from OLS regressions (weighted to adjust for differences in sampling probabilities across sites and over time) of an outcome on indicators for being assigned to the experimental voucher group and the Section 8 voucher group as well as randomization site indicators. Column 4 reports TOT estimates using a 2SLS specification, instrumenting for voucher take-up with the experimental and Section 8 assignment indicators. Standard errors, reported in parentheses, are clustered by family. Panel A restricts the sample to children below age 13 at random assignment; panel B includes children between age 13 and 18 at random assignment. The estimates in panels A and B are obtained from separate regressions. The number of individuals is 2,922 in panel A (except in column 5, where it is 1,625) and 2,331 in panel B. The dependent variable in column 1 is individual W-2 wage earnings, summing over all available W-2 forms. Column 1 includes one observation per individual per year from 2008–2012 in which the individual is 24 or older. Column 2 replicates column 1 using individual earnings as the dependent variable. Individual earnings is defined as the sum of individual W-2 and non-W-2 earnings. Non-W-2 earnings is adjusted gross income minus own and spouse's W-2 earnings, social security and disability benefits, and UI payments, divided by the number of filers on the tax return. Non-W-2 earnings is recoded to zero if negative and is defined as zero for non-filers. Column 3 replicates column 2, controlling for the characteristics listed in online Appendix Table 1A. Column 4 reports TOT estimates corresponding to the ITT estimates in column 2. In column 5, we measure earnings in the year when the individual is 26 years old. In column 6, we measure earnings in 2012, limiting the sample to those 24 or older in 2012. Columns 7–9 replicate column 1 with the following dependent variables: employment (an indicator for having positive W-2 earnings), household income (adjusted gross income plus tax-exempt social security benefits and interest income for those who file tax returns, the sum of W-2 wage earnings, SSDI benefits, and UI benefits for non-filers, and zero for non-filers with no W-2 earnings, SSDI, or UI benefits), and individual earnings growth (the change in individual earnings between year $t − 5$ and the current year t).
 ***Significant at the 1 percent level.
 **Significant at the 5 percent level.
 *Significant at the 10 percent level.

Source: Chetty et al. (2016)

that older treated children were about 4 p.p. less likely to attend college that the control children, according to the ITT estimate. For those who did attend college, they seem to have attended worse colleges, measured by the average earnings of college graduates.

Table 7.4 Treatment effects for economic outcomes among MTO children recipients, separated by age

	College attendance (%) ITT					College quality ($) ITT				
	Age 18–20 (1)	Age 18 (2)	Age 19 (3)	Age 20 (4)	Age 21 (5)	Age 18–20 (6)	Age 18 (7)	Age 19 (8)	Age 20 (9)	Age 21 (10)
Panel A. Children < age 13 at random assignment										
Exp. versus control	2.509** (1.143)	2.213* (1.200)	2.579* (1.452)	2.734* (1.464)	0.409 (1.474)	686.7*** (231.2)	670.2*** (240.6)	800.6*** (274.3)	589.3** (262.3)	337.8 (269.9)
Sec. 8 versus control	0.992 (1.264)	1.221 (1.303)	0.502 (1.613)	1.252 (1.599)	−0.371 (1.592)	632.7** (256.3)	592.0** (268.2)	604.7** (304.7)	701.4** (294.9)	549.2* (293.7)
Observations	15,027	5,009	5,009	5,009	5,009	15,027	5,009	5,009	5,009	5,009
Control group mean	16.5	11.3	18.6	19.6	20.1	20,914.7	20,479.6	21,148.7	21,115.7	21,152.3
Panel B. Children age 13–18 at random assignment										
Exp. versus control	−4.261** (1.712)	−5.866*** (2.180)	−4.460** (2.162)	−2.995 (2.077)	−3.528* (1.972)	−882.8** (385.5)	−1195.7** (482.8)	−890.0* (465.0)	−672.6 (414.2)	−687.9* (402.6)
Sec. 8 versus control	−3.014* (1.785)	−3.339 (2.295)	−3.928* (2.243)	−1.882 (2.182)	−4.455** (2.030)	−597.2 (434.2)	−581.5 (546.9)	−730.2 (511.5)	−492.1 (465.7)	−603.0 (446.6)
Observations	5,100	1,328	1,722	2,050	2,234	5,100	1,328	1,722	2,050	2,234
Control group mean	15.6	12.4	16.8	16.6	17.2	21,638.0	21,337.3	21,880.1	21,629.8	21,597.8

Notes: All columns report ITT estimates from OLS regressions (weighted to adjust for differences in sampling probabilities across sites and over time) of an outcome on indicators for being assigned to the experimental voucher group and the Section 8 voucher group as well as randomization site indicators. Standard errors, reported in parentheses, are clustered by family. Panel A restricts the sample to children below age 13 at random assignment; panel B includes children between age 13 and 18 at random assignment. The estimates in panels A and B are obtained from separate regressions. The dependent variable in column 1 is an indicator for attending college in a given year (having one or more 1098-T tax forms filed on one's behalf), pooling data over the three years when the individual is ages 18–20 with one observation per year per individual. Years before 1999 are excluded because 1098-T data are available beginning only in 1999. Columns 2–5 replicate column 1, using college attendance at each age between 18 and 21 as the dependent variable. The dependent variable in column 6 is Chetty, Friedman, and Rockoff's (2014) earnings-based index of college quality, again pooling data from ages 18–20 starting in 1999. This index is constructed using US population data as the mean earnings at age 31 of students enrolled in that college at age 20; children who do not attend college are assigned the mean earnings at age 31 of children who are not enrolled in any college at age 20. Columns 7–10 replicate column 6, using college quality at each age between 18 and 21 as the dependent variable.
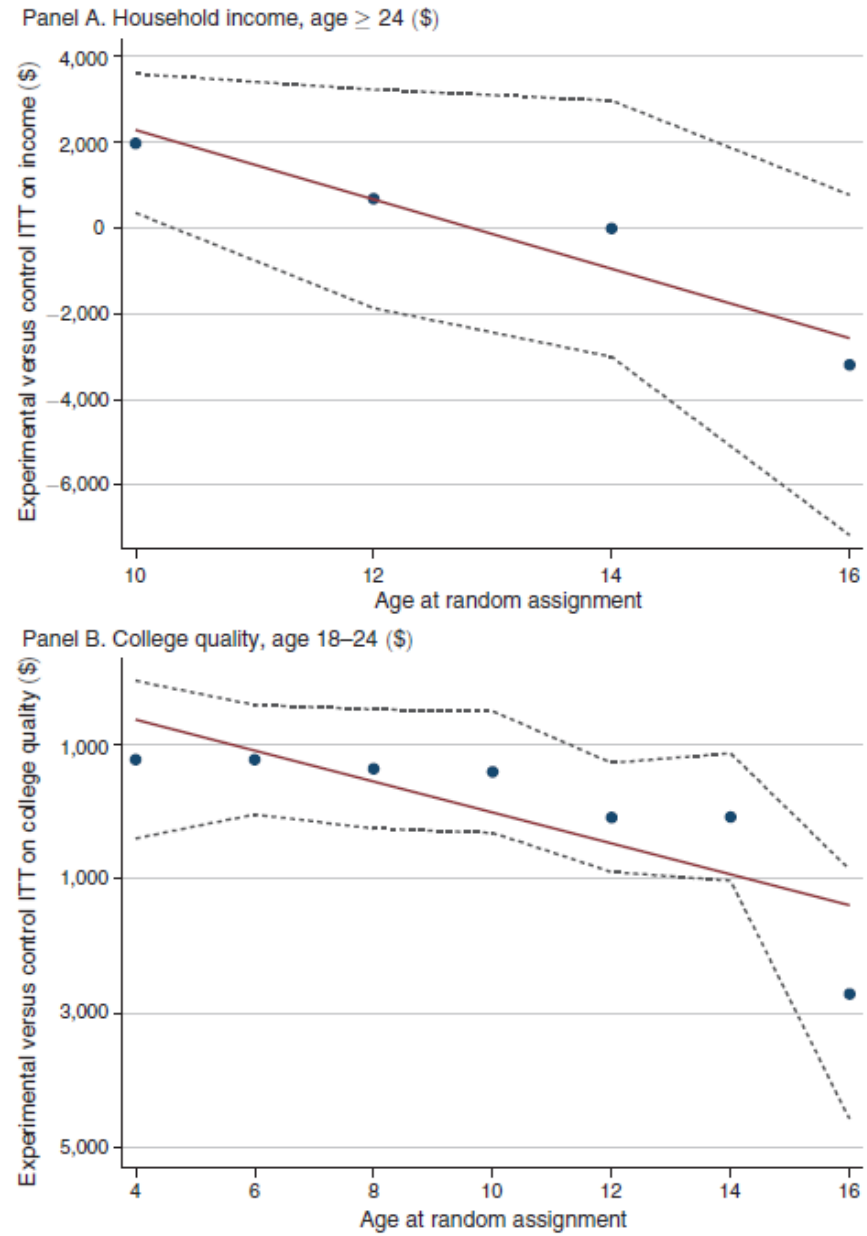***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Source: Chetty et al. (2016)

Chetty et al. (2016) attribute these differences to differences in exposure. Younger children experienced more years in the better neighborhoods, and are in them for most of their teenage years. Figure 7.3 shows that the positive effects of MTO decrease with years of exposure.

Last, Chetty et al. (2016) estimate that the effects for adults were negligible, as shown in figure 7.4. It seems that moving is a good investment for families with children, but does not really improve the outcomes of adults.

Fig. 7.3 Effects of earnings for children over age when treated

Panel A. Household income, age $\geq$ 24 ($)

Panel B. College quality, age 18–24 ($)

Source: Chetty et al. (2016)

Panel A. Cumulative years of exposure to low-poverty neighborhoods
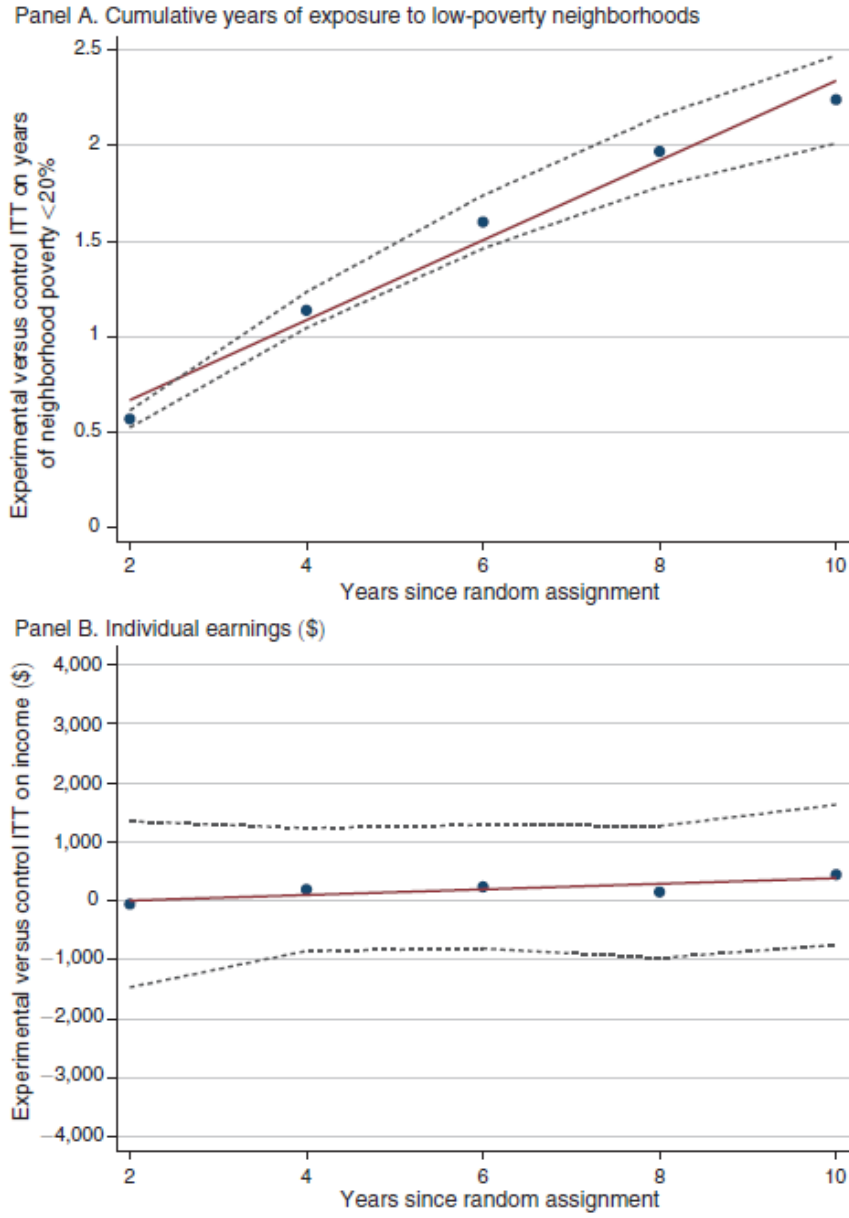

Panel B. Individual earnings ($)

Fig. 7.4

Source: Chetty et al. (2016)

7.2.6 Instrumental variable estimates of neighborhood effects

Although randomized experiments are the gold-standard, we would like to be able to estimate effects in quasi-experimental settings. We will look at Chyn (2018), who uses a novel IV strategy to estimate neighborhood effects.

Before turning to that, we note that a randomized experiment with partial compliance can be interpreted in an IV framework. Let Z be an instrument that affects participation d. Z is a legitimate instrument if it is not directly correlated with outcomes; Z only effects Y by affecting d: $E[Y:0|Z=z] = E[Y_0]$ and $E[Y_1|Z=z] = E[Y_1]$ and participation d is a non-trivial function of Z: $E[d|Z=z]$ is non-trivial function of Z.

If there exist values of Z in the set z_0 that occurs with positive probability and under which $Pr[d_i = 0|Z_i \in z_0] = 1$ and $Pr[d_i = 1|Z_i \notin z_0] > 0$, then one can estimate TOT by defining $e = 1$ for Z not in z_0 and $e = 0$ for Z in Z_0.

In other words, one can estimate the TOT using eligibility (e) as an instrument for treatment (d):

$$2\text{nd Stage}: Y = d\alpha + \eta \tag{7.12}$$

$$1\text{st Stage}: d = e\delta + \mu(1|1) \tag{7.13}$$

$$\widehat{\alpha}_{IV} = \frac{ITT}{P(d=1|e=1)} \tag{7.14}$$

What can one estimate if have a legitimate instrument that affects probability of participation and can be excluded from outcomes equation? The answer is one can estimate a Local Average Treatment Effect (LATE) equal to the average treatment effect on those that can be induced to change their behavior by change in the instrument.

Z_i is a random variable where $P(w) = E[d_i|Z_i = w]$ is a nontrivial function of w. Let treatment for i depend on value of instrument Z: $d_i = d_i(Z_i)$

$$LATE = \alpha_{z,w} = E[Y_{1i} - Y_{0i}|d_i(z) \neq d_i(w)] \tag{7.15}$$

The LATE can be identified if Z is a legitimate instrument (can be excluded from the Y equations) and if have monotonicity condition: $di(z) \geq di(w)$ for all i or $di(z) \leq di(w)$ for all i. Thus we are assuming that there are no "noncompliers" in terms of Angrist-Imbens-Rubin (1996 JASA).

Assume: $d_i(z) > d_i(w)$:

$$E[Y_i|Z_i = z] - E[Y_i|Z_i = w] = (P(z) - P(w)) * E[Y_{1i} - Y_{0i}|d_i(z) - d_i(w) = 1] \tag{7.16}$$

The LATE is consistently estimated by the ratio of the difference in sample mean outcomes for those with values of z and w for the instrument over the difference in the fraction who are treated. Table 7.5 clarifies the differences between the groups.

Chyn (2018) uses demolition of public housing in Chicago as an instrument for moving to a better neighborhood. Public housing buildings were usually in a poor state, but the choice over which to be demolished was random, such that in some

	$d_i(0)$	
	0	1
0	$Y_i(1,0) - Y_i(0,0) = 0$ Never-Taker	$Y_i(1,0) - Y_i(0,1) = -(Y_i(1) - Y_i(0))$ Defier
$d_i(1)$		
1	$Y_i(1,1) - Y_i(0,0) = Y_i(1) - Y_i(0)$ Complier	$Y_i(1,1) - Y_i(0,1) = 0$ Always-Taker

Table 7.5 Causal Effects of Z on Y for Population Units Classified by $d_i(0)$ and $d_i(1)$

cases there could be two neighboring buildings and only one of them was demolished. Figure 7.5 shows that the households who were displaced by demolition did move to better neighborhoods, just as the MTO experimental households did.

Fig. 7.5 Poverty of the destination neighborhood for displaced and not displaced households
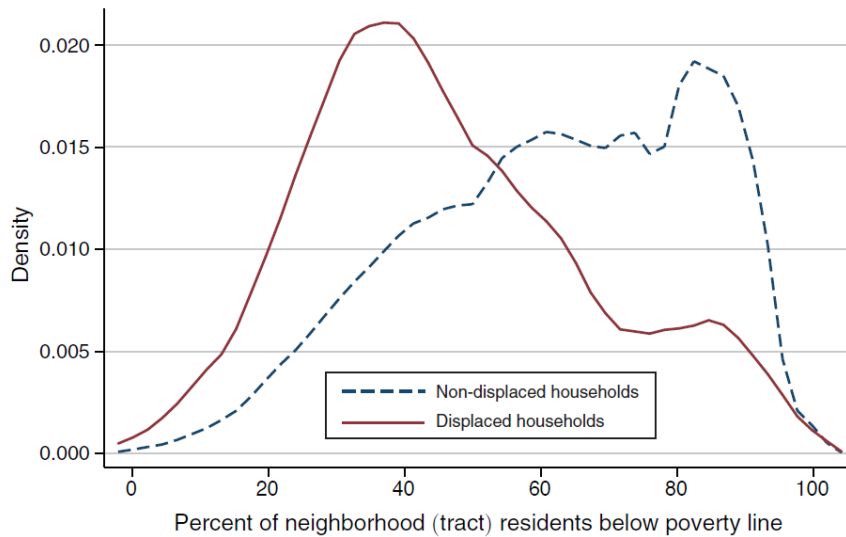
Source: Chyn (2018)

Figure 7.6 shows positive effect of moving in terms of economic and crime outcomes. The results largely confirm those of Chetty et al. (2016), with a couple of exceptions. First, the effects seem to be similar for younger and older children. Second, the effects seem to be larger for women.

Fig. 7.6 Treatment effect estimates of displacement
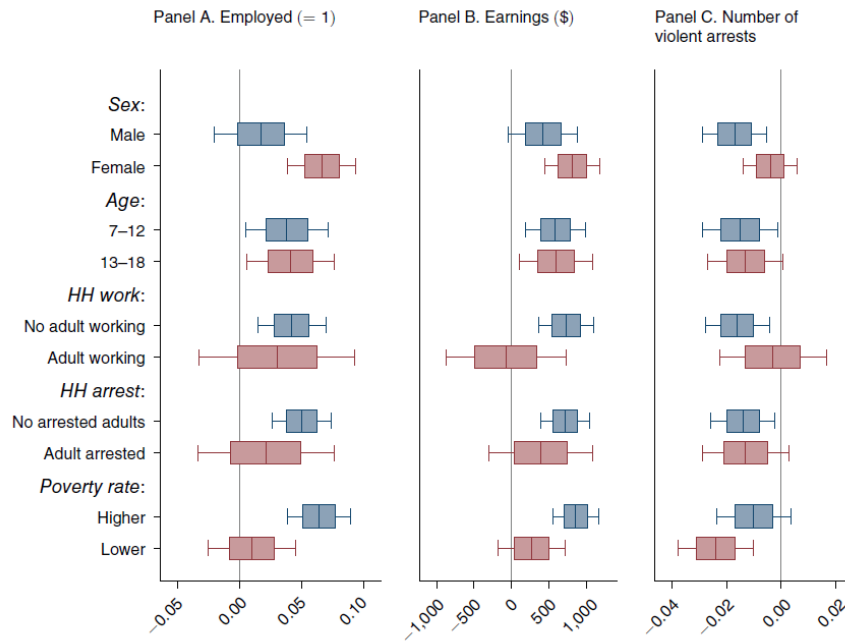
Source: Chyn (2018)

### 7.2.7 Effects on education outcomes

Davis et al. (2017) provide an explanation for why the effects on schooling outcomes were small in the original MTO study. Poverty rates tend to be correlated with value-added in test scores. Value added is just a measure of improvement of test scores at the neighborhood level that controls for children and teacher characteristics.

Low poverty-rate neighborhoods also tend to be more expensive, and it seems that households that come from high poverty rate areas are very sensitive to this. Figure 7.7 shows estimates of sensitivity of utility to rent, from an estimated structural model in LA.

Since households that came from poor areas tended to be more sensitive to rent, when given MTO vouchers, they moved to the cheaper places, which had lower benefits in tends of test scores. Table 7.6 shows a comparison of two policies. One in which MTO recipients choose where to move among low poverty neighborhoods (MTO-A), and one in which they are allocated to a random low poverty neighborhood (MTO-B). When sent to a random neighborhood, children of MTO households see an increase in test scores. But this impact is muted by the sorting of households into the cheaper neighborhoods, so in MTO-A, there is not an impact on test scores.

Fig. 7.7 Sensitivity to rent by neighborhood poverty rate

Figure 9: Average Estimates of α by Tract Poverty Rate
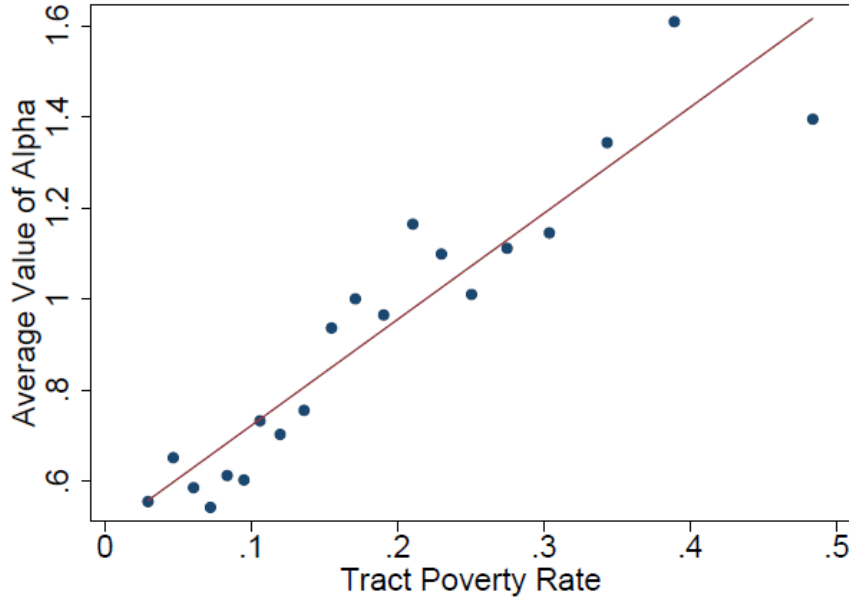
Source: Davis et al. (2018)

Impacts on Woodock-Johnson Math Scores (sd=1)

Exposure time	MTO Demonstration TOT	Simulation Experiments MTO-A (TOT)	MTO-B (ATE of <10% Pov)	p-value H0: (2) = (1)	p-value H0: (3) ≤ (2)
	(1)	(2)	(3)	(4)	(5)
5 years	-0.019	-0.003	0.097	0.919	0.001
10 years	-0.052	-0.008	0.177	0.874	<0.001
18 years	–	-0.002	0.306	–	<0.001

Table 7.6

Source: Davis et al. (2018)

7.3 Taking stock

We have learned about the mechanisms and magnitudes of neighborhood effects. Along the way, we have learned a bit about randomized experiments and IV techniques. The overarching conclusion of these studies is that neighborhood effects are present, although they may only materialize in the long run after long exposure to the better neighborhoods.