Converging Evidence for Neighborhood Effects on Children's Test Scores: An Experimental, Quasi-experimental, and Observational Comparison*

February 24, 2010

Julia Burdick-Will
University of Chicago

Jens Ludwig
University of Chicago and National Bureau of Economic Research

Stephen W. Raudenbush
University of Chicago

Robert J. Sampson
Harvard University

Lisa Sanbonmatsu
National Bureau of Economic Research

Patrick Sharkey
New York University

Corresponding Author:
Jens Ludwig
University of Chicago
1155 East 60th Street
Chicago, IL 60637
(773) 834-0811
jludwig@uchicago.edu

*This paper was prepared for the Brookings Institution "Project on Social Inequality and Educational Disadvantage: New Evidence on How Families, Neighborhoods and Labor Markets Affect Educational Opportunities for American Children," with support from the Russell Sage and Spencer foundations. Thanks to Julian Betts, Rebecca Blank, Greg Duncan, Lawrence Katz, Jeffrey Kling, Richard Murnane, Steven Rivkin, and seminar participants at the Brookings Institution for helpful comments. Thanks also to Ryan Gillette and Matt Sciandra for terrific research assistance. Data for the MTO analyses were provided by the U.S. Department of Housing and Urban Development (HUD). The contents of this chapter are the views of the authors and do not necessarily reflect the views or policies of HUD or the U.S. Government. The authors are listed in alphabetical order; all opinions, and any errors, are of course our own.

I. INTRODUCTION

Educational outcomes vary dramatically across neighborhoods in America. For example, in the Chicago suburb of Wilmette, where the median home value is $441,000, almost everyone graduates from high school and a majority go on to attend, and complete, college. In contrast, the dropout rate is around 44 percent in Chicago's public high schools (Allensworth and Easton, 2001)[1] and is much higher still in some of the city's most disadvantaged neighborhoods. This variation in schooling achievement and other outcomes has generated concern about whether neighborhood environments themselves may influence children's life chances, independent of other individual child and family characteristics. The question is of great importance for public policy in part because residential segregation by income has been increasing since 1970 (Watson, 2009). Despite a decline during the 1990s in the number of people living in some of the most distressed Census tracts (poverty≥40%), a total of 8 million people still lived in such areas in 2000, nearly twice the number as in 1970 (Jargowsky, 2003). Because poor and minority Americans are over-represented in our most disadvantaged neighborhoods, any "neighborhood effects" on children may contribute to persistent disparities in overall schooling outcomes across race and class lines in the U.S.

One reason why neighborhood of residence might affect children's schooling outcomes is through variation across neighborhoods in the quality of local public schools. Another plausible explanation focuses on the social environment. Exposure to more pro-social, higher achieving peers may provide stronger social support for academic achievement, enable children to participate in more developmentally productive study groups, and enable teachers to better target classroom instruction and spend less time dealing with disruptive students. Adults may vary across

1 While there remains some controversy about how to measure school dropout rates, the figure we cite is calculated as the fraction of students enrolled in school at age 13 who go on to drop out of school by age 19. http://ccsr.uchicago.edu/publications/p0a01.pdf

neighborhoods in their capacity and willingness to help monitor local children and enforce community norms, or in their ability to signal the value of staying in school. And exposure to high rates of crime and violence may cause stress, trauma or other mental health problems that negatively affect children's schooling outcomes, might make children more reluctant to go to school or participate in developmentally enriching after-school activities, could hamper the ability of local schools to attract and retain high-quality teachers, and might entice youth to leave school early to earn money in the underground economy (or to join street gangs for protection against criminal victimization). For policy purposes, distinguishing the causal effects of schools from the effects of the social environment is important because in principle the former could be directly addressed by education policy without having to either change the qualities of neighborhood environments outside of schools or else to re-locate low-income families into new neighborhoods.

Empirical claims for the effect of neighborhood context on children's schooling outcomes dates back at least to the Coleman Report, which argued that "attributes of other students account for far more variation in the achievement of minority group children than do any attributes of school facilities and slightly more than do attributes of staff" (Coleman *et al*., 1966, p. 302). These findings, if taken at face value, would seem to imply powerful neighborhood effects on children's learning, given that school composition is determined in large part by neighborhood composition. However, drawing causal inferences from the Coleman Report and most of the subsequent research on peer or neighborhood effects is complicated by the fact that the attributes that lead families to select specific types of neighborhoods may be the same attributes that predict schooling outcomes among children. Because researchers are not always able to capture and control for all of the relevant attributes of a family that influence neighborhood selection, estimates of neighborhood on educational outcomes may be systematically biased. Put differently, educational outcomes could vary across neighborhoods because of the different types of families living in different types of areas, rather than because of any direct causal effects of neighborhood environments on children's outcomes.

The one formal randomized experiment that has been conducted to date to test whether neighborhood environments affect children's life chances is the U.S. Department of Housing and Urban Development's Moving to Opportunity (MTO) residential mobility experiment. MTO has been in operation since 1994 in five cities (Baltimore, Boston, Chicago, Los Angeles, and New York), and has enrolled a total of 4,600 mostly minority public housing families with children. Via random lottery some families but not others were offered the chance to use a housing voucher to relocate to low-poverty census tracts. Random assignment helps solve the selection bias concern with observational studies by generating differences in average neighborhood environments between otherwise similar groups of families, so that any difference we observe in average outcomes across groups can be attributed to the differences in neighborhood trajectories that families experience. Data from the MTO interim study find no statistically significant impacts, on average, on reading or math test scores for children in MTO measured 4-7 years after baseline (Sanbonmatsu et al., 2006). The interim data did show positive and statistically significant effects on reading scores for African Americans in the experimental group. However, due to the number of subgroups examined, it is unclear whether this subgroup effect reflects differential treatment impacts across sub-groups, or reflects sampling variability.

How should one weight the findings from the randomized MTO experiment versus the larger body of non-experimental research, much of which has shown strong neighborhood effects on educational outcomes? This question remains the topic of ongoing discussion (and some disagreement) within the research and policy communities (see for example Clampet-Lundquist and Massey, 2008, Ludwig et al., 2008, and Sampson 2008). Some have interpreted the MTO findings as providing sufficient evidence to conclude that neighborhood environments per se are not very important for children's schooling outcomes. Others have been reluctant to draw this conclusion, in part because of the sizable body of observational and quasi-experimental research suggesting important neighborhood effects and the uncertainty about the practical importance of any selection bias concerns with these studies. Moreover, some critics have expressed skepticism about whether MTO generated sufficiently large changes in neighborhood environments, particularly in racial composition, to adequately test the neighborhood effects hypothesis.

In this paper we try to reconcile the experimental, quasi-experimental and observational research literature regarding neighborhood effects on children, and argue that the available findings may be more convergent than many people believe. Drawing on a number of recent and high-quality quasi-experimental and observational studies, together with a re-examination of MTO findings across the MTO demonstration sites, we believe the evidence allows us to reject the null hypothesis that neighborhood environments *never* matter for children's outcomes. At the same time, the data also do not support the hypothesis that neighborhoods *always* matter. In our view the key question for research and public policy is to learn more about the conditions under which neighborhoods matter for children's academic outcomes, and why—either/or hypotheses are unlikely to capture the complex realities of social life and indeed the data do not support them.

Our ability to answer the questions posed in the present paper is restricted by the limited number of studies that have employed sufficiently strong research designs to support inferences about neighborhood effects on children's outcomes, and by the fact that a disproportionately large share of the studies that meet this research-design threshold have been carried out in a single city, Chicago. With these qualifications in mind, we believe there is at least a suggestive case to be made that children's test scores may be most strongly affected by community violence, or may respond "non-linearly" to levels of concentrated neighborhood disadvantage or community violence. In other words, what may matter most for children's cognitive development is to avoid living in the most severely economically distressed or dangerous neighborhoods in the country, neighborhoods that are found in cities like Baltimore and Chicago but are less prevalent even in other major American cities such as Boston, Los Angeles, and New York. Given the limitations of the available evidence, we offer these as hypotheses to be tested further, rather than strong conclusions.

In the next section of the paper we review literature on neighborhood effects on children's academic outcomes, focusing mostly on the studies that employ strong research designs. The third section of the paper explores candidate explanations for why neighborhood environments might matter more for some children in some circumstances than others. The fourth section discusses potential implications of our hypotheses about what features of neighborhood environments might be most relevant for children's academic outcomes.

II. NEIGHBORHOOD EFFECTS AND EDUCATION-RELATED OUTCOMES

The causal effects of different aspects of the neighborhood environment on schooling outcomes among children remains a subject of disagreement (Dietz, 2002; Ellen and Turner, 1997; Leventhal and Brooks-Gunn, 2000; Sampson, Morenoff and Gannon-Rowley, 2002). Families choose their places of residence under more or less severe constraints, given family size, income, local housing prices, and varied levels of racial or other kinds of discrimination in the housing market. As a result of these differential constraints and family preferences, family characteristics are systematically associated with place of residence. Disentangling the causal effects of neighborhood environments from those difficult-to-measure attributes that may be relevant for both residential selection and the key behavioral outcomes of interest is a major challenge for this empirical literature. In light of these methodological concerns, and given the large number of good reviews of the "neighborhood effects" literature that are already in circulation, we provide here a more selective discussion of particularly influential national studies of neighborhood effects on educational outcomes, before focusing in on a more recent set of unusually strong observational or quasi-experimental studies, which happen to have been all carried out in Chicago; and the five-city HUD-funded Moving to Opportunity (MTO) randomized housing mobility experiment.

A. Neighborhood correlations on schooling outcomes

Two studies conducted with data from the Panel Study of Income Dynamics (PSID) provide evidence for strong neighborhood effects on children's test scores and schooling outcomes, while a third using the same dataset, but different methods, finds null effects. Harding (2003) uses the PSID data to compare outcomes of children who are matched with respect to their family background characteristics but who are living in different types of neighborhoods, and finds strong evidence for important neighborhood effects. For blacks and non-blacks there are only very slight differences in school dropout rates between youth living in low poverty tracts (<10% poor) versus moderate poverty rates (10-20% poor) during adolescence. On the other hand he finds large differences in dropout rates between those living in low versus high (>20% poor) census tracts during adolescence; these effects are equal to around twelve percentage points for both blacks and non-blacks, which are very large compared to the baseline dropout rate of 20-25 percent for youth in low-poverty tracts. Similarly, in a highly influential early study of neighborhood effects Brooks-Gunn and her colleagues (1993) find evidence that it is the absence of affluent adults within a census tract, rather than the presence of disadvantaged neighbors, that is most strongly predictive of children's test scores (see also Brooks-Gunn, Duncan and Aber, 1997a,b).

The results from these studies contrast with the findings from Plotnick and Hoffman (1999), who use sibling fixed effects to study neighborhood effects on educational attainment, among other outcomes. Using variation in neighborhood characteristics among sisters in the PSID to estimate the effects of neighborhoods on the probability of receiving postsecondary education, the results show null effects for each measure of neighborhood disadvantage examined. This is one of several studies that have questioned the presence of neighborhood effects on methodological grounds, an issue we return to below.

B. Observational and Quasi-Experimental Findings from Chicago

To date there have been four major studies of neighborhood effects on children's schooling outcomes carried out in Chicago. Three of the four studies find evidence of large gains in children's academic outcomes from living in less rather than more distressed neighborhoods. The one study that yields contradictory findings follows a sample of families who are involuntarily displaced by public housing demolitions, which raises the possibility that whether families benefit from living in less distressed areas may depend on whether they want to live in such areas.

Perhaps the most extensive observational study of neighborhood effects that has been carried out to date is the Project on Human Development in Chicago Neighborhoods (PHDCN), which followed a racially and socio-economically mixed sample of children ages 0-18 and living in Chicago as of 1995 (see Table 1). A random sample of about 6,000 children and their primary caregivers were interviewed in 1995-7, and then again in 1997-9 and in 1999-2002.

Three of us (Sampson, Sharkey and Raudenbush, 2008) analyzed verbal cognitive ability[2] among African American children living in neighborhoods that vary with respect to an index of neighborhood concentrated disadvantage. This index is a weighted average of six census tract

2 This verbal cognitive ability measure is a composite of the Wechsler Intelligence Scale for Children vocabulary test and the Wide Range Achievement Test reading examination. Math ability was not assessed in the PHDCN.

characteristics: share of residents who receive welfare; tract share poor; share unemployed; share female-headed households; share African American; and share of the tract under 18 years old. The analysis compares outcomes for African American children living in census tracts that fall in the top quarter of the concentrated disadvantage distribution (N=237, average concentrated disadvantage index value of 2.52) versus the rest of the sample (N=543, average index value of 1.58). A key strength of the PHDCN design is that it follows children over time, meaning the analysis can control for past residence within a high or low disadvantage neighborhood. The estimated effect of living in concentrated disadvantage is driven in large part by comparing the outcomes of children who stay in such neighborhoods over time with those of other children who move from very disadvantaged into less disadvantaged areas, or vice versa (that is, children who begin in low disadvantage neighborhoods and move to high disadvantage neighborhoods).[3]

The analysis suggests living in the most disadvantaged quarter of Chicago neighborhoods (statistically, this is associated with roughly a one standard deviation difference in the scale of concentrated disadvantage) reduces children's verbal test scores by around one quarter of a standard deviation (see Figure 1). This "effect size" is roughly equivalent to missing one or two years of schooling. There is also some evidence of an age interaction, such that the influence of concentrated disadvantage may be greatest for younger children (Sampson 2008).

---

3 All of the time-varying covariates in the PHDCN analysis are controlled using inverse probability of treatment weighting (IPTW) as introduced by Robins, Hernan, and Brumback (2000) and extended to the multilevel setting by Hong and Raudenbush (2008). Like related forms of propensity score matching, inverse probability of treatment may be susceptible to bias from unobserved characteristics, but it has the advantage over standard least squares regression of being less sensitive to assumptions about the functional form of the relationship between the observable covariates and the outcomes of interest. IPTW gives relatively low weight in the analysis to people who receive the "treatment" (concentrated neighborhood disadvantage) that they are predicted to have a very high likelihood of receiving. Previous residence in a concentrated disadvantage neighborhood is a very strong predictor for future residence in such a neighborhood, so that the observations that receive the largest weights in the analysis are those children who were living in concentrated disadvantaged neighborhoods in a previous wave of the PHDCN but move to a less distressed neighborhood in a subsequent wave of the survey, or vice versa.

While the PHDCN findings have been influential within social science, at least as important for housing policy have been the findings from the Gautreaux mobility program in Chicago. This program was named after the plaintiff (Dorothy Gautreaux) in a 1966 racial discrimination lawsuit filed against the Chicago Housing Authority (CHA) and HUD. The lawsuit charged discrimination based on the heavy concentration of African American families in public housing projects located in high-poverty areas. The U.S. Supreme Court agreed in 1976 and ordered CHA to provide housing vouchers to African American public housing residents that could be used only in neighborhoods in the city or suburbs that were less than 30 percent Black. Units were assigned to eligible families on a waiting list of approximately 2000 families a year (Rubinowitz and Rosenbaum 2000). Some of those apartments were in areas of Chicago that were poor and segregated, but improving, while others were able to get apartments in low-poverty, predominantly white or integrated suburban areas (Mendenhall, DeLuca, and Duncan 2006).

A 1988 follow-up survey of 342 families who used Gautreaux vouchers found that moving to the Chicago suburbs versus other parts of the city was associated with significant improvements in young adult's later educational attainment. Compared to the surveyed students who remained in the city of Chicago, suburban movers were four times less likely to have dropped out of school (20 percent versus 5 percent); more likely to be in a college track in high school (24 versus 40 percent); twice as likely to attend any college (21 percent versus 54 percent); and almost seven times as likely to attend a four-year college (4 percent versus 27 percent). The only educational attainment measure for which the suburban students did not appear to be doing significantly better than the city students was their grade point average, which could reflect higher grading standards in suburban schools (Rubinowitz and Rosenbaum 2000: 134-6).

While the Gautreaux program has been extremely influential, the study was nevertheless not a true randomized experiment. Families may have had some choice in whether or not they accepted the first apartment offered to them, and indeed there is some evidence that the baseline characteristics of families that ended up in the suburbs are systematically different from those who ended up in the city (Mendenhall, Duncan and Deluca, 2006, Votruba and Kling, 2009). This has made researchers nervous that the Gautreaux city and suburban movers may have been different with respect to pre-existing *unobserved* characteristics as well, which could lead analysts to confound the causal effects of suburban moves with the influence of these unmeasured attributes that may affect outcomes as well as the likelihood of moving to the suburbs.

However a more recent experimental study of Chicago's housing voucher program, which relies on true random assignment of families to different neighborhood environments, seems to support the basic conclusion from Gautreaux (Ludwig et al., 2010). In July 1997 the private firm running the city's voucher program, called CHAC, Inc., opened the city's housing voucher program wait list for the first time in a dozen years. A total of 82,607 income-eligible households, almost all of whom were black (see Table 1 and, for more details, Appendix Table 1), applied and were then randomly assigned to the program wait list. Starting in August 2007 the families were offered vouchers in order of their wait-list position. Roughly 4,625 families were offered vouchers in the first year of the program, and by May 2003 around 18,110 families had been offered housing vouchers, at which point CHAC was over-leased and stopped offering vouchers.

Ludwig et al. (2010) focus on families who were living in public housing at the time they applied to CHAC for a voucher; the analytic sample is composed of children who are 4-11 at baseline. Families who received a voucher experienced changes in neighborhood environments that are fairly similar to those observed among MTO families, a point that we discuss in more detail below. These voucher–supported moves increase children's achievement test scores in reading and math on the Iowa Test of Basic Skills (ITBS). The effect of being offered a housing voucher, known in the program evaluation literature as the "intent to treat" (ITT) effect, was equal to around .05 and .08 standard deviations for reading and math scores, respectively. Since only around one-quarter of CHAC families with children relocated using a voucher, the effects of actually leasing up with a voucher (the "effects of treatment on the treated," or TOT) effects of voucher receipt are on the order of .2 and .3 standard deviations for reading and math, respectively (see Figures 1 and 2).

Jacob (2004) uses variation in neighborhood conditions generated by the demolition of public housing in Chicago and finds little systematic evidence of any achievement test score changes among children. His analytic sample consists of around 10,500 mostly African American children living in Chicago public housing in the mid-1990s, when the Chicago Housing Authority began to demolish housing projects with federal funding. Jacob argues that the timing of which projects were demolished first is driven by random events at the projects (e.g., broken pipes, etc.) Public housing demolitions lead children to move into census tracts with poverty rates that are about 15 percentage points lower than those of children who stay in public housing (who have an average tract poverty rate of 68 percent). Yet the difference in reading and math scores on the Iowa tests for children who do versus do not move is less than .01 standard deviations. The 95% confidence interval around this estimate enables Jacob to rule out impacts that are any larger than about .05 standard deviations (Figures 1 and 2).

One candidate explanation for why the children in Jacob's sample do not show the same gains in test scores as children in the other three Chicago studies noted here could be that only families who want to live in less distressed areas may benefit. Data from the MTO study discussed next reveal that only around one-quarter of eligible public housing volunteered for that mobility program (Goering et al., 2003, p. 11), which suggests that a majority – perhaps a large majority – of families who are displaced by public housing demolitions may not have wanted to move. Of course in any comparison of results across such a small number of studies, alternative explanations for differences in study findings are also possible.

C. Moving to Opportunity

Motivated by the suggestive findings of Chicago's Gautreaux mobility program, in the early 1990s HUD decided to fund a large-scale randomized housing-mobility experiment known as Moving to Opportunity (MTO). Eligibility for MTO was limited to families living at baseline in public housing in selected high-poverty census tracts in five U.S. cities (Baltimore, Boston, Chicago, Los Angeles, and New York City). Starting in 1994 HUD began randomly assigning eligible low-income families with young children who volunteered to participate in MTO into three different groups: the *experimental group* was offered a housing voucher that could only be used in neighborhoods where the poverty rate was 10 percent or less according to the 1990 census, and was given relocation counseling assistance; the *"Section 8" housing voucher group* were offered standard housing vouchers that could be used for any unit that met basic standards, but were not restricted geographically; and a *control group*, that did not receive any special MTO funding, but could receive any of the regularly available social services for which they would have been eligible.

In total, 4,600 families signed up between 1994 and 1997 to be randomly assigned to one of the three groups. Of households assigned to the MTO experimental group, 47 percent used an MTO voucher to relocate to a low-poverty Census tract, while 62 percent assigned to the regular Section 8 housing voucher group relocated through MTO. Compliance rates vary across MTO cities.

Data from the interim MTO study found no overall statistically significant impacts on either broad reading or broad math scores on the Woodcock-Johnson-Revised tests measured 4-7 years after baseline (Sanbonmatsu et al., 2006). While data on risky behaviors, delinquency and other youth outcomes found sharp gender differences in MTO impacts (Kling, Ludwig and Katz, 2005, Kling, Liebman and Katz, 2007), with MTO moves generating beneficial changes in outcomes among female youth and on balance adverse behavioral changes among male youth, for test scores we observe no statistically significant changes for either boys or girls (Sanbonmatsu et al., 2006). We also do not observe any statistically significant differences in MTO impacts on test scores by age overall, although it is important to note that the youngest MTO children who were under age 6 at baseline were only beginning their school years.

However when we look separately at Chicago combined with the one other almost entirely black demonstration site in MTO, Baltimore, we do see some evidence of impacts of neighborhood changes on children's achievement test scores. Researchers are usually (and appropriately) cautious about estimating too many sub-group effects because of concerns about "false positives" – if one generates estimates for, say, 20 independent sub-groups, one would expect to see an estimated effect for at least one subgroup that is statistically significant at the usual 5 percent level purely by chance. But there is an important substantive justification for looking separately by site in MTO given the evidence noted above for neighborhood effects in Chicago (at least from three of the four Chicago studies). When we generate separate estimates for the set of African American children enrolled in the Baltimore and Chicago demonstration sites the treatment on the treated effect is equal to .3 standard deviations in reading, with mixed results in math (Figures 1 and 2).[4] The fact

4 The reading and math achievement levels of MTO participants were measured in 2002 by Abt Associates using the Woodcock-Johnson Revised (WJ-R) instrument. The WJR W scores have been adjusted for interviewer effects (see Sanbonmatsu et al. (2006), Appendix 1 for additional details). We estimate the effect on treatment compliers ("TOT"

that the impacts are more pronounced in reading than math is itself interesting, since most studies of school-based interventions tend to find larger impacts on math than on reading.

III. UNDERSTANDING VARIATION IN NEIGHBORHOOD EFFECTS ON CHILDREN

The research literature summarized in the previous section enables us to reject the null hypothesis that "neighborhoods never matter." But the mixed pattern of findings across studies seems to also allow us to reject the alternative null hypothesis that neighborhoods *always* matter. In our view, a key question for both social science and public policy is why and for whom neighborhood environments seem to matter for academic outcomes, and what the implications of that variation in treatment effects might be for policy efforts designed to improve the life chances of some of our nation's most disadvantaged children.

In this section we try to narrow down the set of candidate explanations for the variation documented above in findings about when and how neighborhoods affect children's reading or math achievement test scores. Our ability to convincingly determine which explanations are most important is limited by the small number of very strong study findings and the even smaller number of study sites and independent data samples from which results have been generated.

Before presenting new analysis that attempts to adjudicate among different potential explanations, we begin by highlighting some basic evidence that runs counter to several plausible candidate explanations for why results might vary across studies. Variation in findings does not seem to rest with methodological problems such as selection bias with observational studies, since we find support for neighborhood effects on children's test scores even in studies that use random assignment of families to different mobility conditions. Study results do not seem likely to vary

effect") using a two-stage least squares regression controlling for a series of baseline covariates (see Orr et al. 2003 page B-15 for a complete list). MTO children ranged from roughly age 6 to age 20 at time of testing.

because of slight differences in the age of the samples being studied, since there is considerable overlap in the age of the different study samples, and because age differences in responses to neighborhood environments are not large enough to explain the differences across studies. An alternative candidate explanation is that just a few particularly distressed public housing projects might be responsible for all the findings of neighborhood effects on children's test scores, yet the PHDCN provides evidence for neighborhood effects among a sample that includes few public housing families. It would be surprising if the private-market housing in which these families lived was of lower quality than the public housing in which control families were living in the Boston, Los Angeles and New York demonstration sites. A more subtle hypothesis is that differences in findings within MTO could stem from variation across sites in how the experiment was carried out, but this explanation also does not seem to fit the data very well.[5]

In what follows we show that the variation across studies in findings does not seem to be due to differences in neighborhood effects on children across race or ethnic groups, nor to differences across studies in the size of the changes that children experience with respect to potentially key neighborhood attributes, including local school quality, racial composition and concentrated disadvantage more generally. The evidence we produce does *not* allow us to rule out the possibility that there may be "non-linearities" in the relationship between concentrated neighborhood disadvantage and children's academic outcomes, meaning the effect of a given unit change in neighborhood disadvantage may be greater for children whose starting position is a relatively more disadvantaged neighborhood environment. We also cannot rule out the possibility

5 We might worry that in those demonstration sites where the housing search assistance was least effective, only the most motivated families would relocate as part of the MTO treatment. If more motivated families benefited more from changing neighborhoods, then differences across sites in the composition of who moves through MTO could explain differences across cities in the size of the estimated effects of treatment on the treated for children's test scores. But there is no clear relationship between the size of the MTO impact on test scores and the MTO treatment compliance rates. Relative to the other MTO cities, Baltimore has a relatively high compliance rate among experimental group families (57%) while Chicago has a relatively low experimental-group compliance rate (34%). There is also no clear relationship between the size of the impacts and responses rates by site or by treatment group within site.

that different study samples experience differently-sized changes in neighborhood violence rates, or that there are non-linearities in the relationship between children's test scores and exposure to violence in the community.

A. Differences in vulnerability across demographic groups

One candidate explanation for the apparent discrepancy in results across studies is differences in study populations. Table 1 shows that all of the Chicago samples are almost entirely African American, as is the set of MTO families in the Baltimore site (the one other MTO city besides Chicago where we find evidence that moves to less-distressed areas increase children's test scores). Put differently, some of the strongest empirical evidence for neighborhood effects on children's test scores comes from studies of African American samples, who might be more vulnerable to neighborhood influences than Hispanic or white families, perhaps because of higher rates of single-parent households.[6] In the MTO study, the proportion of "never married" adults at baseline was higher in Baltimore and Chicago (68.5 to 72.7%) than in the other three sites (54.5 to 57.2%). Two-parent households may mitigate any adverse influences of living in a distressed neighborhood by providing more parental supervision, which could reduce exposure to neighborhood influences (if more supervised children are subject to earlier or stricter curfews) or ameliorate adverse neighborhood influences by, for example, intervening at the first sign of trouble and providing academic or social supports. Note that although African Americans in Chicago and Baltimore look similar with respect to marital status compared to African Americans at the other MTO sites, there are differences on baseline characteristics suggesting African Americans in Baltimore and Chicago were slightly more disadvantaged than African Americans at the other 3 sites in terms of work status, teen parenting, and welfare receipt (see also Appendix Table 1).

[6] In a nationally representative sample of kindergarten students conducted in 1998, 15 percent of white students, 50 percent of black students and 24 percent of Hispanic students were living in single-parent households (Duncan and Magnuson, 2005).

We can test and reject this hypothesis with the MTO data, by pooling data from the three MTO sites where there is racial and ethnic diversity in the program populations (Boston, LA and NYC) and examining whether there is evidence for MTO effects on test scores for African Americans in these cities but not for Hispanic children. We find that there is no evidence for test score gains for either blacks or Hispanics in these three MTO cities. A different way to test this hypothesis is to use the PHDCN data, which (unlike samples studied in the Chicago MTO site, the Chicago CHAC voucher study, or the Chicago public housing demolition study) does sample Hispanics as well as African Americans. Within the Chicago PHDCN data, we find at least suggestive evidence that living in a disadvantaged neighborhood may have adverse impacts on the verbal scores for Hispanic as well as African American children (see Appendix Figure 1).

B. Local School Quality

Since most public schools draw their students from the local community, it is plausible that much or even most of the variation across neighborhoods that exists in children's test scores could in principle be due to variation across areas in school quality. Neighborhoods might vary in the quality of their local schools because of differences in political power in securing resources from centralized public school bureaucracies, or because many teachers tend to prefer to teach schools serving more affluent student bodies (Hanushek, Rivkin and Kain, 2004, Hanushek and Rivkin, 2007), which might make it hard to recruit and retain the best teachers in high-poverty areas.

One hypothesis to explain differences across study findings, then, is differences in the degree to which variation in neighborhood environments is associated with the underlying quality of the schools that children attend. Attention to this hypothesis is motivated in part by the fact that Gautreaux is perceived to have generated large changes in school quality (since suburban schools are thought to be so much better than Chicago public schools), while Table 2 shows that children who move to less distressed areas through MTO still attend struggling schools. For example in the full MTO sample, the average child assigned to the MTO treatment group attends a school where only around 25 percent of students are at or above national norms, which is higher than the control group average by around 8 percentage points but still suggests these children are attending fairly low-performing schools overall. The minimal change in school quality induced by MTO led Dobbie and Fryer (2009, p. 22) to conclude that "a better community, as measured by poverty rate, does not significantly raise test scores if school quality remains essentially unchanged."

Table 2 shows that children in MTO's Chicago and Baltimore samples experience larger changes in school racial and class composition than do children in Boston, LA and NYC, but do not seem to experience larger gains in the one measure of "school quality" available – the percent of students in a school who are at or above national norms on reading and math tests.[7] In the Chicago CHAC voucher study, children who move to a less distressed area have higher test scores than the control group despite the fact that they do not experience any gain in the share of children scoring above national norms. Recognizing the limitations of this school quality measure, the results in Table 2 taken at face value would nevertheless seem to argue against the hypothesis by Dobbie and Fryer (2009) that test scores are unresponsive to changes in neighborhood environments absent changes in school quality.

C. Neighborhood racial composition

7 Note that this story would not be likely to change much if we used our data to calculate some sort of school "value added" measure that adjusted school-wide average test scores for the socio-demographic composition of the school's student body.If the MTO mobility treatment caused children to move into schools serving relatively more disadvantaged student bodies compared to the schools of the control group children, then we might worry that what looks like a fairly modest difference between the average treatment versus control school in the share of children meeting national norms might actually reflect large differences in underlying school "value added" to student learning, since the treatment schools in this case would be achieving slightly better average student outcomes among a more disadvantaged student population. But there is unlikely to be any hidden value-added advantage to the school serving MTO treatment group children since these schools are serving slightly less disadvantaged student bodies compared to control group schools and achieving only slightly better average student outcomes.

A different hypothesis for the variation in impacts on children's test scores comes from differences across studies in the change that families experienced in neighborhood racial segregation. Perhaps most famously, the Gautreaux mobility program in Chicago was required to move families into racially mixed neighborhoods. In contrast, MTO focused on moving families into lower-poverty areas, which it did, but MTO did not induce major changes in neighborhood racial composition among participating families (Table 3; see also Appendix Table 3).

Some have argued that the lack of change in neighborhood racial segregation induced by MTO undermines the study's capacity to provide a rigorous test of the neighborhood effects hypothesis since racial composition might itself be a crucial aspect of a child's neighborhood (Clampet-Lundquist and Massey, 2008).[8] However, Table 3 shows that families in the MTO Chicago and Baltimore sites do not experience much larger changes in neighborhood racial segregation, despite starting in neighborhoods with much higher concentrations of African Americans, than do families in the other three MTO sites where children did not experience any gains in achievement test scores as a result of their MTO moves. In other words, racial composition was not much affected anywhere by the MTO experiment, so it cannot explain site differences. Table 3 also shows families in the Chicago CHAC voucher study experience no declines in the share of their census tract that is black relative to controls (or in the share of the census tract that is minority, broadly defined), and yet these moves are still sufficient to increase children's achievement test scores. While we do not have a great many data points, the available

8 As Clampet-Lundquist and Massey (2008, pp. 115-6) argue: "Because of the history of segregation and continuing barriers to realizing residential preferences … relative to areas inhabited by middle-class whites, Asians, or Latinos, those inhabited by the black middle class exhibit lower property values, higher crime rates, lower employment rates, higher levels of unwed childbearing, poorer schools, lower educational achievement, and higher rates of welfare dependency … Even though middle-class black areas may not themselves display concentrated poverty, because of racial segregation they tend to be located adjacent to or very near areas of concentrated deprivation and often share common service catchment areas."

evidence seems to suggest that changes in neighborhood racial composition are not *necessary* for improved educational outcomes and do not explain the divergent findings across sites.

Concentrated neighborhood disadvantage

While differences across studies and samples in the size of the changes in neighborhood racial segregation children experience do not seem to explain variation across studies in children's achievement test score gains, what about other aspects of neighborhood disadvantage? What Table 1 from above suggests is that almost all of the best empirical evidence to date for neighborhood effects on children's learning comes from studying African American families living in neighborhoods that are much more disadvantaged than what we see in other cities. The next-to-last row of Table 1 shows for each of our study samples the values of the concentrated disadvantage index used by Sampson, Sharkey and Raudenbush (2008), which is a weighted average of poverty, percent black, percent adults unemployed, percent households with a female head, percent residents on welfare, and percent of residents under age 18 (see Appendix Table 1 for details). We focus on the concentrated disadvantage index in order to have a consistent measure of neighborhood environments across different studies. We note that the disadvantage index has a strong negative correlation with the presence of affluent neighbors, which is the neighborhood measure that seems particularly predictive of youth outcomes in the analyses by Brooks-Gunn and her colleagues (1993).[9]

The mean value of the neighborhood concentrated disadvantage index for the public housing families in the new Chicago housing voucher sample studied by Ludwig et al. (2010) was 3.39, and equaled 3.16 for the Chicago MTO sample and 2.74 for the pooled samples of families in

9 In Chicago, our concentrated disadvantage index has a correlation of -.83 with the share of families in the tract with incomes of at least $30,000; using national data, the correlation is quite similar at -.82.

the Baltimore and Chicago MTO sites together. By comparison the average value of the concentrated disadvantage index in the three other MTO sites (Boston, Los Angeles, and New York) was equal to just 1.51, and was 2.2 for the African American PHDCN sample. Much, but certainly not all, of the difference in concentrated disadvantage in Baltimore and Chicago compared to the other three MTO cities is due to the substantially greater level of racial segregation in Chicago and Baltimore. This can be seen in the last row in the table, when we re-calculate the concentrated neighborhood disadvantage, but now exclude the measure of tract percentage black from its construction.

Table 3 shows that those samples that experience the largest changes in achievement test scores do not experience unusually large changes in concentrated disadvantage. Figures 1 and 2 showed us that achievement test score gains are largest for families in the Baltimore and Chicago MTO sites, for families in the new Chicago housing voucher study, and for African American families in the PHDCN. The changes in the concentrated neighborhood disadvantage scale experienced by families in these three samples equal -.399, -.548, and -.935, respectively (second to last row, Table 3). At least the first two of these numbers are not substantially different from what we see in the three MTO sites (Boston, LA and NYC) where there are no detectable test score impacts (-.530). The fact that Baltimore and Chicago seem to be different from our other study samples with respect to baseline concentrated disadvantage *levels* but not *changes* leads naturally to a hypothesis that neighborhood effects on children's outcomes may be non-linear. The web appendix discusses several statistical tests that we have carried out to formally test for non-linearities. While our analyses do not yield clear, convincing evidence for such non-linearities, it is important to note that our tests have relatively weak statistical power.

E. Exposure to Community Violent Crime

In addition to the possibility of non-linearities between concentrated neighborhood disadvantage and children's test scores, another candidate explanation for variation in impacts across studies that we cannot reject is exposure to community violence. The two MTO cities in which we find evidence for neighborhood effects on children's outcomes, Baltimore and Chicago, have greatly elevated levels of crime and violence compared to the other three MTO cities. For example 1998 homicide rates per 100,000 equaled 47.1 in Baltimore and 25.6 in Chicago, compared to 6.1 in Boston, 11.8 in Los Angeles and 8.6 in New York (see the web appendix for additional details). This raises the possibility that, like concentrated disadvantage, there may be a non-linear relationship between exposure to extremely violent neighborhood settings and children's test scores. Unfortunately it is even more difficult to test for non-linearity in the effects of crime or violence because of data limitations and the associated difficulty in making comparisons of crime data across cities. There is considerable variation across cities in the areal level at which crime rates are available: the data for Baltimore are from 9 police "beats", whereas the data for Boston are from 11, compared to 18 in Los Angeles, 76 in New York City, and 279 in Chicago.[10]

With these differences in mind, it is possible to examine how change in exposure to area-level crime rates relates to changes in test scores. Figure 3 plots the averages of both measures separately for each MTO site and randomized mobility group (experimental, Section 8, and control).[11] For each data point we have subtracted off the overall mean level of beat-level violence in that MTO site, since all of the statistical analyses of MTO data always compare the average outcomes of the randomized mobility groups within sites (that is, control for site fixed effects). The

10 In some cities these police department administrative units are districts or areas instead of beats, although for convenience we refer to all of these areas as "beats" since what we mean is the smallest geographic area for which we are able to obtain crime data for the cities and years that are relevant for the MTO study.
11 See Web Appendix 4 for raw score differences.

line fit through these data points in the figure shows the correlation between beat-level violence and children's test scores, and shows there is a negative relationship between beat violent crime and children's test scores, which is larger for reading than for math, and that this relationship is driven by the Baltimore and Chicago sites (as seen by the regression lines that are fitted by dropping data from those two MTO sites).[12]

Sharkey's (2009) analysis of data from the PHDCN provides some confirming support for the violence-achievement link using variation over time across Chicago neighborhoods. He compares the outcomes of children in the PHDCN within the same neighborhood who were interviewed and tested at different points in time, and finds that African American children interviewed within a week of a homicide occurring in their neighborhood had achievement test scores around one-half standard deviations lower than other children, suggesting a large acute effect of violence on achievement scores. Because these analyses come from comparing outcomes for children living in the same neighborhood (that is, from models that control for neighborhood fixed effects), the results are not simply picking up the fact that test scores are generally lower in some neighborhoods than others within the city of Chicago.[13] Additional support comes from Grogger's (1997) analysis of High School and Beyond, which suggests high school graduation rates are lower in schools in which principals report more serious problems with crime and violence.

IV. CONCLUSION

---

12 Fitting a regression line through these MTO site-group means is essentially equivalent to generating instrumental variables estimates for the relationship between beat-level violent crime and children's test scores using interactions of indicators for MTO treatment group assignment and MTO site as instruments for local violent crime (Kling, Liebman and Katz, 2007, Ludwig and Kling, 2007).

13 We have also tried to carry out other within-city analyses by examining whether children in the Chicago CHAC voucher study who live in more violent baseline neighborhoods (and so presumably experience the largest changes in beat-level violent crime) experience the largest test score changes. Unfortunately these results are not very informative because they are relatively imprecisely estimated. For example for a one standard deviation change in beat-level violent crime rates (around 68 per 10,000) we could not rule out a relationship that is as large (in absolute value) as around -.15 standard deviations in reading or math achievement test scores.

Most of the empirical evidence supporting neighborhood effects on children's educational outcomes came from observational studies such as the PHDCN, that follow families wherever they wind up living, or quasi-experimental studies of government mobility programs such as Gautreaux. In contrast, the one randomized mobility study, HUD's Moving to Opportunity (MTO) demonstration, found no detectable evidence on children's achievement test scores, on average, across the five program sites (Baltimore, Boston, Chicago, Los Angeles, and New York).

This has led to a variety of different hypotheses that seek to reconcile the apparently conflicting evidence. Some researchers conclude that the most important feature of neighborhood environments for children's learning must be racial segregation, since Gautreaux and PHDCN compare families who live in neighborhoods with different levels of racial as well as economic segregation, while MTO generates large changes in economic segregation but limited changes in racial segregation. Other researchers hypothesize that children may be unresponsive to changes in neighborhood environments after a certain age, since many of the MTO children examined in the 5 year follow up were already of school age at baseline when their families relocated. A third hypothesis stems from the observation that MTO generated relatively little change in school characteristics, and suggests the possibility that neighborhood environments simply may not matter very much for children's achievement test scores on their own.

Our re-examination of the available data plus results from a new housing voucher lottery in Chicago lead us to reject the null hypothesis that neighborhood environments are irrelevant for children's achievement test scores. Our reading of the evidence suggests that changing neighborhoods can improve children's achievement test scores even without changes in neighborhood racial segregation or school quality, and that even children who have already spent many years living in segregated, economically distressed and dangerous neighborhoods can experience gains in cognitive outcomes from moving.

But moving to a less distressed neighborhood does not inevitably produce this outcome – treatment-by-city interactions seem to be an important part of the story. Namely we believe we can conclude that moves to less distressed areas in Chicago and Baltimore improve children's test scores while that does not appear to be the case in the other three MTO sites of Boston, Los Angeles, and New York. This does not mean neighborhood effects in general are not present in these latter cities, of course. It may be that the kinds of changes in neighborhood environments that MTO induced in these cities do not significantly change educational test scores, but do impact other outcomes.

One candidate explanation for this pattern is non-linearity in the relationship between neighborhood concentrated disadvantage and children's achievement test scores, a hypothesis bolstered by the observation that baseline levels of concentrated disadvantage are much higher in Baltimore and Chicago than in the other three cities we examine here (even though the "treatment dose" in terms of the change in neighborhood attributes is not systematically different across cities). Our argument for that explanation is mostly circumstantial, since our direct tests of the non-linearity hypothesis have relatively low statistical power. We can assemble a similar (but still circumstantial) argument that exposure to violence explains variation in children's academic achievement – Baltimore and Chicago have much higher rates of violence than the other three MTO cities, and the MTO treatment group assignment generates larger changes in exposure to violence in Baltimore and Chicago as well, although from much higher levels to begin with. In addition, within the PHDCN study, there is evidence that test scores are lower for children tested within a week of a homicide occurring in their neighborhood compared to children from the same neighborhood assessed at a further point in time from the most recent local homicide.

The evidence presented here will hopefully provoke a change in the conversation around neighborhood effects on children's learning, and refocus attention away from a narrow examination of the role of schools and neighborhood racial segregation and toward a broader examination of why neighborhood influences might vary across cities. Of particular importance is the possibility that neighborhood effects on children are non-linear or may be related to community violence.

Were future research to support the importance of these two mechanisms, one potential implication for public policy would be to focus scarce housing policy resources on trying to de-concentrate the most severely disadvantaged neighborhoods in the U.S., of the sort found in places like Baltimore and Chicago, but less so in other major American cities such as Boston, Los Angeles and New York. The existence of non-linear relationships between concentrated disadvantage and children's outcomes suggests that re-sorting poor children across neighborhoods would lead to an increase in overall average achievement, consistent with Guryan's (2004) finding that court-ordered school desegregation starting in the late 1960s led to declines in black dropout rates with no detectable changes in schooling outcomes for whites. If more were known about the specific aspects of concentrated disadvantage that were most helpful to children, then in principle community-development strategies or mixed-income housing as well as residential mobility strategies could be employed to help improve the life chances of poor children in these areas. Such policies might also require subsidies to either non-poor families to live alongside poor families (in mixed-income developments), or subsidies to poor families to make relatively longer-distance moves into higher-poverty neighborhoods. In any case evidence for non-linearities in neighborhood effects would suggest the great importance of prioritizing scarce housing policy resources on the most distressed areas, given that at present just 28 percent of income-eligible poor families receive assistance under existing means-tested housing programs (Olsen, 2003).

Were community violence to be confirmed as a key contributor to children's cognitive development, one implication might be that policymakers interested in children's cognitive development and success in school should expand their focus outside of the school setting, and consider policies relating to the provision of effective policing and the provision of safe community environments for children. Shifts in police practices or increased policing, done well, may achieve short-term changes in the developmental quality of some of our nation's most disadvantaged neighborhoods (Sherman, 2003, Evans and Owens, 2007). Donohue and Ludwig (2007) argue that each additional dollar spent on policing generates from four to six dollars in benefits to society just from the increase in well being of community residents, setting aside the possibility of any developmental benefits to children. Interventions designed to provide safe and enriching environments for children, both within and outside the schools, also represent promising policy options. The Harlem Children's Zone is the best known example of a program that has attempted to provide a "conveyor belt" of services that would enhance the environment for an entire community of children, but only the school component of the program has been evaluated to this point (Dobbie and Fryer 2009; see also Tough, 2008). But perhaps the main point, and a key theme underlying this entire volume, is that some promising ways to improve children's schooling outcomes may have little at all to do with schools.

REFERENCES

Allensworth, Elaine, and John Q. Easton (2001) "Calculating a Cohort Dropout Rate for the Chicago Public Schools: A Technical Research Report," Consortium on Chicago School Research.

Brooks-Gunn, Jeanne, Greg J. Duncan, Pamela Kato Klebanov, and Naomi Sealand (1993) "Do neighborhoods influence child and adolescent development?" *American Journal of Sociology*. 99(2): 353-395.

Brooks-Gunn, Jeanne, Greg J. Duncan, and J. Lawrence Aber, Editors (1997a) *Neighborhood Poverty, Volume 1: Context and Consequences for Children*. New York: Russell Sage.

Brooks-Gunn, Jeanne, Greg J. Duncan, and J. Lawrence Aber, Editors (1997b) *Neighborhood Poverty, Volume 2: Policy Implications in Studying Neighborhoods*. New York: Russell Sage.

Clampet-Lundquist, Susan and Douglas S. Massey (2008) "Neighborhood effects on self-sufficiency: A reconsideration of the Moving to Opportunity experiment," *American Journal of Sociology*. 114(1): 107-143.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York (1966) "Equality of Educational Opportunity," Washington, D.C.: Government Printing Office.

Dietz, Robert D. (2002) "The estimation of neighborhood effects in the social sciences: An interdisciplinary approach," *Social Science Research.* 31 (4): 539-575.

Donohue, John J., and Jens Ludwig (2007) "More COPS," Policy Brief #158, Brookings Institution.

Dobbie, Will and Roland G. Fryer, Jr. (2009) "Are high quality schools enough to close the achievement gap? Evidence from a bold social experiment in Harlem," Working Paper, Harvard University Department of Economics (April 2009 draft).

Duncan, Greg, and Katherine Magnuson (2005) "Can Family Socioeconomic Resources Account for Racial and Ethnic Test Score Gaps?" *The Future of Children*. (1): 35-54.

Ellen, Ingrid Gould, and Margery Austin Turner (1997) "Does neighborhood matter?. Assessing recent evidence," *Housing Policy Debate*. 8(4).

Evans, William, and Emily Owens (2007) "COPS and Crime," *Journal of Public Economics*. 91(1-2): 181-201.

Goering, John, Judith D. Feins, and Todd M. Richardson (2003) "What have we learned about housing mobility and poverty deconcentration?" In *Choosing a Better Life? Evaluating the*

Moving to Opportunity Social Experiment. Edited by John Goering and Judith D. Feins. Washington, DC: Urban Institute Press. pp. 3-36.

Grogger, Jeffrey T. (1997) “Local violence and educational attainment,” *Journal of Human Resources*. 32(4): 659-682.

Guryan, Jonathan (2004). “Desegregation and Black Dropout Rates,” *American Economic Review*. 94(4): 919–943.

Hanushek, Eric A., John F. Kain, and Steve G. Rivkin (2004) “Why public schools lose teachers,” *Journal of Human Resources*. 39(2).

Hanushek, Eric A. and Steve G. Rivkin (2007) “Pay, working conditions and teacher quality.” *The Future of Children*. Spring 2007.

Harding, David J. (2003) “Counterfactual models of neighborhood effects: The effect of neighborhood poverty on dropping out and teenage pregnancy.” *American Journal of Sociology*. 109(3): 676-719.

Jacob, Brian A. (2004) “Public housing, housing vouchers and student achievement: Evidence from public housing demolitions in Chicago,” *American Economic Review*. 94(1): 233-258.

Jargowsky, P. A. (2003). *Stunning progress, hidden problems: The dramatic decline of concentrated poverty in the 1990s*. Living cities census series. Washington, D.C.: Brookings Institution.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz (2007) "Experimental Analysis of Neighborhood Effects," *Econometrica.* 75(1):83-119.

Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz (2005) “Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment,” *Quarterly Journal of Economics*. 120(1). 87-130.

Leventhal T, and J. Brooks-Gunn (2000) "The neighborhoods they live in: the effects of neighborhood residence on child and adolescent outcomes," *Psychological Bulletin*. 126 (2): 309-37.

Ludwig, Jens and Jeffrey Kling (2007) “Is Crime Contagious?” *Journal of Law and Economics*. 50(3): 491-518.

Ludwig, Jens, Jeffrey Liebman, Jeffrey Kling, Greg J. Duncan, Lawrence F. Katz, Ronald C. Kessler,and Lisa Sanbonmatsu (2008) “What can we learn about neighborhood effects from the Moving to Opportunity experiment? A comment on Clampet-Lundquist and Massey,” *American Journal of Sociology*. 114(1):144-88.

Ludwig, Jens, Brian A. Jacob, Michael Johnson, Greg J. Duncan and James E. Rosenbaum (2009) "Neighborhood effects on low-income families: Evidence from a randomized housing voucher lottery," Working Paper, University of Chicago.

MacIntyre, Sally and Anne Ellaway (2003) "Neighborhoods and Health: An Overview." Pp. 20-42 in *Neighborhoods and Health*, edited by Ichiro Kawachi and Lisa Berkman. Oxford: Oxford University Press.

Mendenhall, Ruby, Stefanie DeLuca, and Greg Duncan (2006) "Neighborhood resources and economic mobility: Results from the Gautreaux Program," *Social Science Research*. 892-923.

Olsen, Edgar O. (2003) "Housing Programs for Low-Income Households." In *Means-Tested Transfer Programs in the United States*. Edited by Robert A. Moffitt. Chicago: University of Chicago Press. pp. 365-442.

Plotnick, Robert D. and Saul Hoffman (1999) "The Effect of Neighborhood Characteristics on Young Adult Outcomes: Alternative Estimates," *Social Science Quarterly*. 80:1-8.

Rubinowitz, Leonard S. and James E. Rosenbaum (2000) *Crossing the Class and Color Lines*. Chicago: University of Chicago Press.

Sampson, Robert J. (2008) "Moving to Inequality: Neighborhood effects and experiments meet social structure," *American Journal of Sociology*. 114(1): 189-231.

Sampson, Robert J., Jeffrey D. Morenoff, and Thomas Gannon-Rowley (2002) "Assessing 'Neighborhood Effects': Social processes and new directions in research," *Annual Review of Sociology*. 28: 443-478.

Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls (1997) "Neighborhoods and violent crime: A multi-level study of collective efficacy," *Science*. 27:918-924.

Sampson, Robert J., Patrick Sharkey, and Stephen W. Raudenbush (2008) "Durable effects of concentrated disadvantage on verbal ability among African American children." *Proceedings of the National Academy of Sciences*. 105: 845-853.

Sanbonmatsu, Lisa, Jeffrey R. Kling, Greg J. Duncan, and Jeanne Brooks-Gunn (2006) "Neighborhoods and Academic Achievement: Results from the MTO Experiment," *Journal of Human Resources.* 41(4):649-691.

Sharkey, Patrick (2009) "The acute effect of local homicides on children's cognitive performance," Working Paper, New York University.

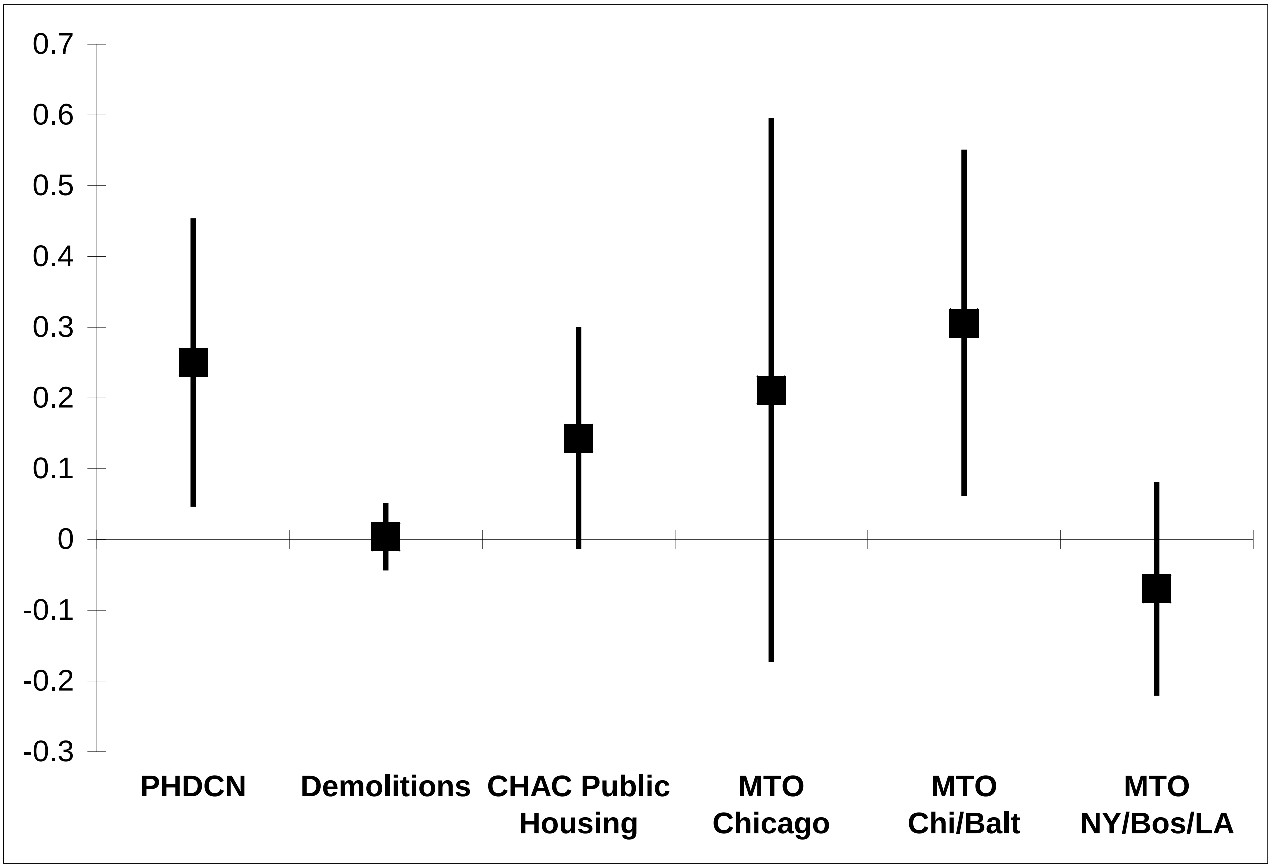
Sherman, Lawrence W. (2003) Editor of "Misleading Evidence and Evidence-Led Policy: Making Social Science More Experimental." *Annals of the American Academy of Political and Social Science.* 589:6-19.

Tough, Paul (2008) *Whatever It Takes: Geoffrey Canada's Quest to Change Harlem and America*. New York: Houghton Mifflin Harcourt.

Votruba, Mark Edward and Jeffrey R. Kling (2009) "Effects of neighborhood characteristics on the mortality of black male youth: Evidence from Gautreaux, Chicago," *Social Science and Medicine*. 68(5): 814-823.

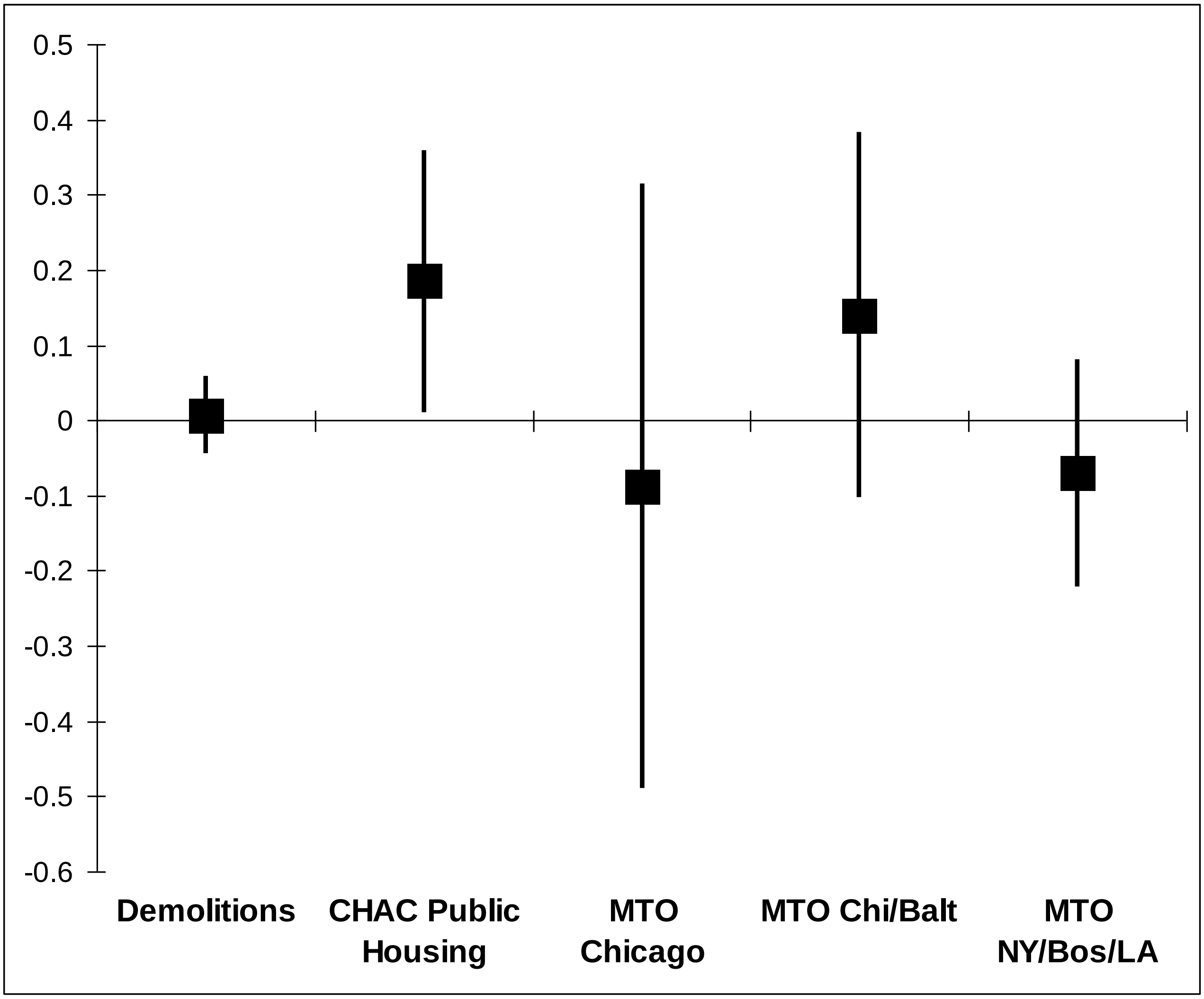
Watson, Tara (2009) "Inequality and the Measurement of Residential Segregation by Income in American Neighborhoods." *Review of Income and Wealth*. 55(3): 820-844.

Figure 1: Summary of Effects of Different Studies on Children's Verbal Test Scores



Notes: The x-axis lists the name of each study: Project on Human Development in Chicago Neighborhoods (Sampson, Sharkey and Raudenbush, 2008); Chicago public housing demolition study (Jacob 2004); Chicago CHAC voucher study for families living in public housing at baseline (Ludwig et al., 2010); and results from the Moving to Opportunity (MTO) study for different cities (Sanbonmatsu et al., 2006). The y-axis shows the estimated effect of changing neighborhoods on children's verbal test scores in each of the studies, expressed as an effect size (share of a standard deviation in the test score distribution, so that an effect size of .2 means children living in less distressed areas have average scores about one-fifth of a standard deviation higher than children living in more distressed areas). For the mobility studies we are presenting effects of actually moving through the program (the effects of treatment on the treated, or TOT).

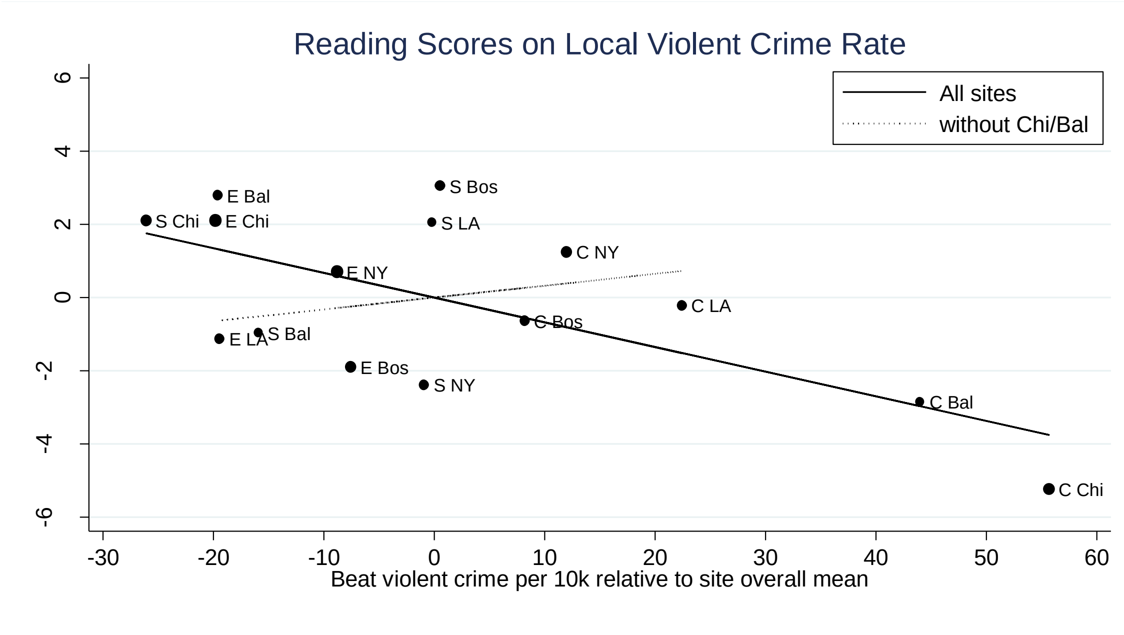
Figure 2: Summary of Effects of Different Studies on Children's Math Test Scores

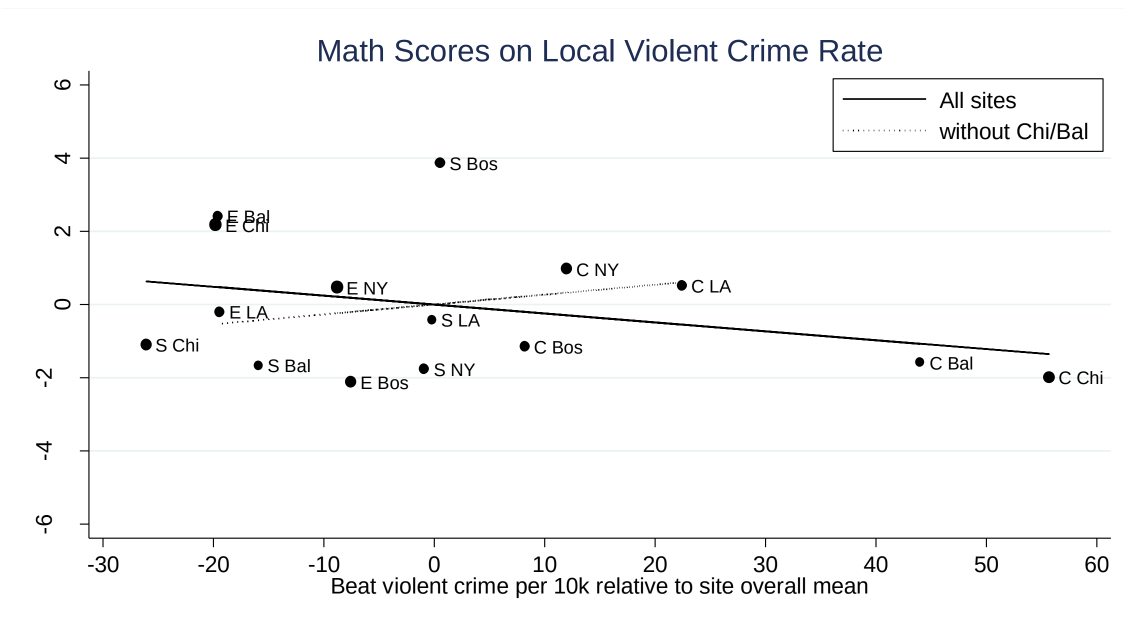


Notes: The x-axis lists the name of each study: Project on Human Development in Chicago Neighborhoods (Sampson, Sharkey and Raudenbush, 2008); Chicago public housing demolition study (Jacob 2004); Chicago CHAC voucher study for families living in public housing at baseline (Ludwig et al., 2010); and results from the Moving to Opportunity (MTO) study for different cities (Sanbonmatsu et al., 2006). The y-axis shows the estimated effect of changing neighborhoods on children's math test scores in each of the studies, expressed as an effect size (share of a standard deviation in the test score distribution, so that an effect size of .2 means children living in less distressed areas have average scores about one-fifth of a standard deviation higher than children living in more distressed areas). For the mobility studies we are presenting effects of actually moving through the program (the effects of treatment on the treated, or TOT).

Figure 3

Relationship between beat-level violent crime and children's test scores Across MTO Demonstration Cities and Randomized Mobility Groups





Notes: The figures plot the average beat- or district-level violent crime rate (x-axis) and average Woodcock-Johnson-Revised reading score (top panel) or math score (bottom panel) for MTO families broken out by whether families were assigned to the MTO experimental, Section 8-only, or control groups, and by site (Baltimore, Boston, Chicago, Los Angeles, and New York City). We re-scale each group's test score and beat violent crime rate by subtracting off the average values for test scores and beat violent crime rates within that MTO site. The solid lines in each figure show the correlation between beat violent crime rates and test scores implied by the 15 data points (i.e. the regression line fit through these points), while the dashed line in each figure shows what happens to this relationship when we drop the data points for the Baltimore and Chicago sites.

Table 1: Comparing Baseline Characteristics Across Study Samples

	Gautreaux	Public Housing Demolitions	PHDCN: African American	PHDCN: Hispanic	CHAC: Public Housing	CHAC: in MTO Tract at Baseline	MTO: Full Sample	MTO: Chicago Only	MTO: Chicago, Baltimore Only	MTO: NY, LA, Boston
Child Age	8.47	10.34	9.01	8.93	7.76	7.67				
		(4.01)	(2.52)	(2.49)	(2.21)	(2.25)				
Household Head Characteristics:										
Age	36.06		36.83	35.34	30.51	30.05	34.09	32.49	32.91	34.81
			(9.30)	(6.93)	(6.64)	(6.33)	(9.08)	(8.78)	(8.78)	(9.18)
African American	1.00	1.00	0.98	0.01	0.98	0.99	.67	.99	.99	.47
			(0.13)	(0.09)	(0.13)	(0.08)	(.4)	(.09)	(.12)	(.50)
Hispanic			0.00	0.95	0.01	0.00	.29	.01	.01	.46
			(0.04)	(0.22)	(0.09)	(0.03)	(.45)	(.08)	(.11)	(.50)
Employed			0.53	0.52	0.35	0.33	.27	.27	.26	.27
			(0.50)	(0.50)	(0.48)	(0.47)	(.43)	(.43)	(.43)	(.44)
Receiving Welfare	50.03		0.48	0.23	0.83	0.85	.74	.81	.81	.71
			(0.50)	(0.42)	(0.38)	(0.36)	(.43)	(.39)	(.39)	(.45)
Neighborhood Characteristics:										
Tract Poverty Rate		0.84	0.27	0.22	0.61	0.71	.50	.66	.58	.45
		(0.11)	(0.13)	(0.10)	(0.19)	(0.11)	(.14)	(.10)	(.15)	(.12)
Tract share black			0.76	0.13	0.89	0.99	.59	.99	.90	.39
			(0.29)	(0.18)	(0.24)	(0.06)	(.33)	(.04)	(.23)	(.21)
Concentrated disadvantage index			2.20	0.70	3.00	3.39	2.18	3.16	2.74	1.84
			(1.11)	(0.85)	(0.77)	(0.33)	(.72)	(.29)	(.71)	(.46)
Concentrated disadvantage index (without % black)			1.93	0.84	2.25	2.56	1.69	2.34	1.99	1.51
			(1.18)	(0.87)	(0.61)	(0.31)	(.51)	(.27)	(.55)	(.38)

Notes: This table reports baseline household and neighborhood characteristics for the different studies that we review: Project on Human Development in Chicago Neighborhoods (Sampson, Sharkey and Raudenbush, 2008); Chicago public housing demolition study (Jacob 2004); Chicago CHAC voucher study for families living in public housing at baseline (Ludwig et al., 2010); and results from the Moving to Opportunity (MTO) study for different cities (Sanbonmatsu et al., 2006). The concentrated disadvantage index is a weighted average of several different census tract-level characteristics including tract share poor, tract share black, tract share unemployed, tract share households headed by a female, tract share on welfare, and share of the tract's population that is under age 18.

Table 2: Control Means and Effects of Voucher-Assisted Residential Mobility at Follow-Up On Average School Characteristics

	CHAC: Public Housing at Baseline	CHAC: in MTO Census Tract at Baseline	MTO: Full Sample	MTO: Chicago Only	MTO: Chicago and Baltimore Only	MTO: NY, LA and Boston
Proportion Black						
Control Mean	.899	.954	0.557	0.914	0.902	0.343
Voucher movers vs. non-movers	-.048*	-.022	-0.048*	-0.086	-0.098*	-0.031
	(.025)	(.027)	(.023)	(.063)	(.042)	(.023)
Proportion Hispanic						
Control Mean	.075	.031	0.307	0.042	0.029	0.479
Voucher movers vs. non-movers	.034*	.009	-0.054*	0.015	0.004	-0.077*
	(.020)	(.016)	(.017)	(.036)	(.020)	(.023)
Proportion Receiving Free Lunch						
Control Mean	.929	.936	0.726	NA	0.699	0.733
Voucher movers vs. non-movers	-0.037***	-.035***	-0.093*	NA	-0.191*	-0.071*
	(.008)	(.010)	(.020)	NA	(.042)	(.022)
Proportion At/Above National Norms						
Control Mean	.304	.282	0.169	0.104	0.128	0.194
Voucher movers vs. non-movers	-.021	.014	0.077*	.083*	0.068*	0.086*
	(.013)	(.021)	(.018)	(.039)	(.029)	(.022)

Notes: This table reports the effects of relocating using a housing voucher on different school characteristics reported at left; that is, each cell in the table represents the difference in average school characteristics for children who moved with a voucher versus the average for those children in the control group who would have moved had their families been assigned a voucher (the effect of treatment on the treated, or TOT). The voucher effect cells report the difference in average characteristics with the standard error underneath reported in parentheses. Each column reports results for a different study and/or sample within a study: Project on Human Development in Chicago Neighborhoods (Sampson, Sharkey and Raudenbush, 2008); Chicago public housing demolition study (Jacob 2004); Chicago CHAC voucher study for families living in public housing at baseline (Ludwig et al., 2010); and results from the Moving to Opportunity (MTO) study for different cities (Sanbonmatsu et al., 2006).

Table 3: Control Means and Effects of Voucher-Assisted Mobility at Follow-Up – Neighborhood Characteristics

	CHAC: Public Housing at Baseline	CHAC: in MTO Census Tract at Baseline	MTO: Full Sample	MTO: Chicago Only	MTO: Chicago and Baltimore Only	MTO: NY, LA and Boston
Tract Poverty Rate						
Control Mean	0.481	0.467	0.392	0.419	0.387	0.394
Voucher movers vs. non-movers	-0.274***	-0.336	-0.190*	-0.187*	-0.141*	-0.213*
	(0.094)	(0.259)	(.019)	(.069)	(.041)	(.018)
Tract share black						
Control Mean	0.837	0.912	0.548	0.857	0.848	0.371
Voucher movers vs. non-movers	0.028	-0.112	-0.023	0.04	-0.059	-0.01
	(0.091)	(0.287)	(.028)	(.089)	(.057)	(.029)
Concentrated disadvantage index						
Control Mean	2.057	2.170	1.869	2.307	2.192	1.678
Voucher movers vs. non-movers	-0.548**	-1.012	-0.488*	-0.411	-0.399*	-0.530*
	(0.258)	(0.809)	(.067)	(.242)	(.145)	(.064)
Concentrated disadvantage index (without % black)						
Control Mean	1.357	1.408	1.409	1.59	1.482	1.366
Voucher movers vs. non-movers	-0.572***	-0.918	-0.466*	-0.444	-0.350*	-0.516*
	(0.215)	(0.648)	(.052)	(.19)	(.111)	(.051)

Notes: This table reports the effects of relocating using a housing voucher on different neighborhood characteristics reported at left; that is, each cell in the table represents the difference in average neighborhood characteristics for children who moved with a voucher versus the average for those children in the control group who would have moved had their families been assigned a voucher (the effect of treatment on the treated, or TOT). The voucher effect cells report the difference in average characteristics with the standard error underneath reported in parentheses. Each column reports results for a different study and/or sample within a study: Project on Human Development in Chicago Neighborhoods (Sampson, Sharkey and Raudenbush, 2008); Chicago public housing demolition study (Jacob 2004); Chicago CHAC voucher study for families living in public housing at baseline (Ludwig et al., 2010); and results from the Moving to Opportunity (MTO) study for different cities (Sanbonmatsu et al., 2006). The concentrated disadvantage index is a weighted average of several different census tract-level characteristics including tract share poor, tract share black, tract share unemployed, tract share households headed by a female, tract share on welfare, and share of the tract's population that is under age 18.

WEB APPENDIX

Converging Evidence for Neighborhood Effects on Children's Test Scores: An Experimental, Quasi-experimental, and Observational Comparison*

February 10, 2010

Julia Burdick-Will
University of Chicago

Jens Ludwig
University of Chicago and National Bureau of Economic Research

Stephen W. Raudenbush
University of Chicago

Robert J. Sampson
Harvard University

Lisa Sanbonmatsu
National Bureau of Economic Research

Patrick Sharkey
New York University

This web appendix presents additional material for our chapter on the relationship between neighborhood environments and children's academic achievement.

Appendix Table 1 compares the average baseline child, household and neighborhood characteristics for people in the different neighborhood effect studies that we review in our chapter.

Appendix Table 2 reports the average neighborhood conditions during the study period for the control group in each of our studies, as well as the effects of voucher-assisted mobility on these average neighborhood characteristics (the effects of treatment on the treated).

Appendix Table 3 reports average crime rates for each of the five cities in the Moving to Opportunity study sample for a selected year (1998). MTO enrolled and randomly assigned families over the period from 1994-98; relative rankings of MTO cities with respect to their levels of homicide or other crimes are similar if we look at crime data from other years.

Appendix Figure 1 reports the results of testing for neighborhood effects on verbal scores in the Project on Human Development in Chicago Neighborhoods (PHDCN) sample separately for African American versus Hispanic children. In the originally published PHDCN paper Sampson, Sharkey and Raudenbush (2008) treated extreme concentrated disadvantage as a binary variable and therefore discarded children who had no probability of ever living in such extreme conditions. In this analysis, however, we treat concentrated disadvantage as a continuous variable and are able to use data from all of the Hispanics (N=733) and African Americans (N=1066) in the original sample. More specifically, we use a propensity score matching model to use covariates from the first wave of the longitudinal PHDCN samples to predict the level of concentrated neighborhood disadvantage that families will experience during the second wave of the study. We then stratify the sample on these predicted "dosages," with 23 strata for blacks and 25 strata for Hispanics. Within strata there is balance in baseline covariates, although at the same time there is also some variation within these strata in the neighborhood environments that families actually experience. Under the assumption that within strata, variation in actual neighborhood environments is uncorrelated with other determinants of children's learning, then we can fit a model to the data that weights the different strata-specific relationships in concentrated disadvantage and verbal scores to get an overall estimated relationship. The best model for both Hispanics and African Americans is a negatively sloped line. While the negative relationship is not quite statistically different from zero for Hispanics, at the same time we also cannot reject the null hypothesis that this negative relationship is the same as the one we find for African Americans.

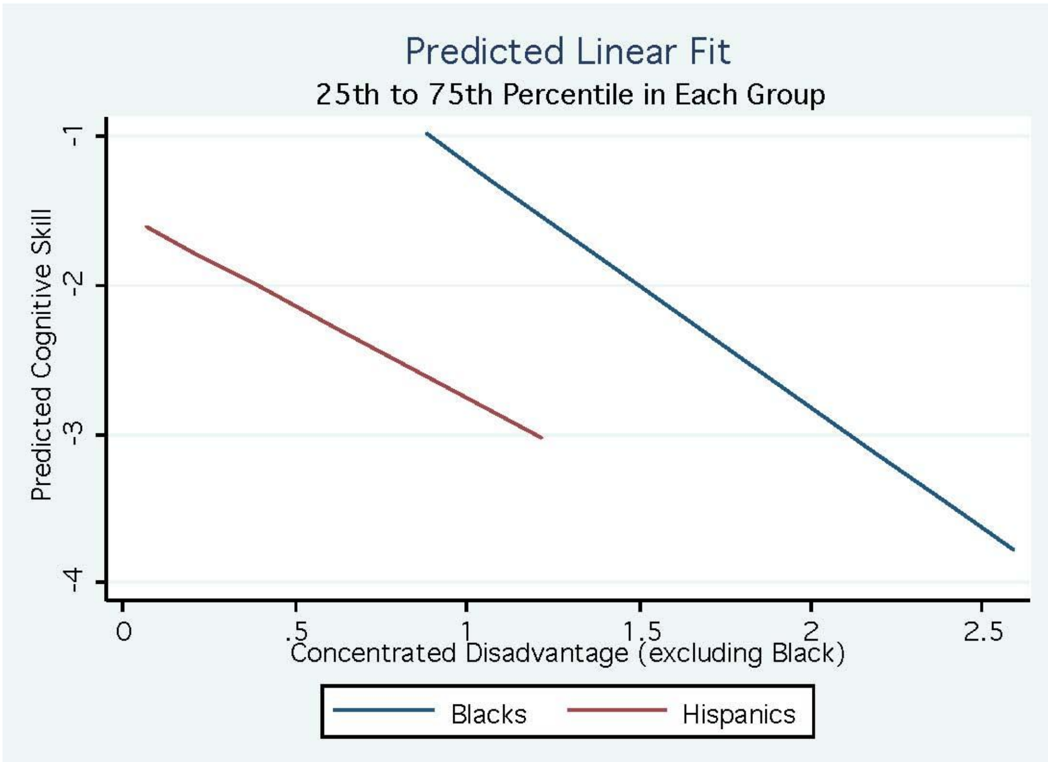
Appendix Figures 2 and 3 report the results of testing for non-linearity of concentrated neighborhood disadvantage on children's verbal scores using data from MTO, and provide one way to explore the possibility of non-linearities visually. We plot the level of the average concentrated disadvantage index and average reading and math scores (converted to Z-scores with mean of 0, standard deviation of 1) for the compliers and would-be compliers in each MTO treatment and control group in each site, and then connect these points, so that the horizontal distance covered by each line shows the treatment-on-the-treated effect in that site on concentrated disadvantage while the vertical distance covered by each line shows the treatment-on-the-treated effect on test scores. Appendix Figure 2 provides some suggestive evidence for non-linearity in reading scores, although this is less clear for math in Appendix Figure 3. Of course with just 15 data points our ability to draw strong conclusions is quite limited.

We have also carried out a series of empirical analyses to formally test for non-linearity in the relationship between neighborhood concentrated disadvantage and children's test scores. Our first test of non-linearity comes from the PHDCN data. As noted above, our analysis calculates a predicted neighborhood "dosage" for all African Americans and Hispanics in the dataset and then stratifies the sample with respect to this predicted dosage to ensure balance in baseline covariates for people who actually experience different neighborhood environments and so for whom we would wish to compare outcomes. Within each of the 23 strata for African Americans and 25 strata for Hispanics, we first examine whether the best-fitting function between concentrated disadvantage and children's verbal scores is linear or quadratic. In no case can we reject linearity, although of course within strata the range of concentrated disadvantage that families experience is limited so it is perhaps not surprising that a straight line serves as a reasonable local approximation. Our other test is to fit a two-level model to the data and see whether there is variation across strata in the slopes of these lines. While we cannot reject the null that all the slopes are the same, this is not a very high-powered test.

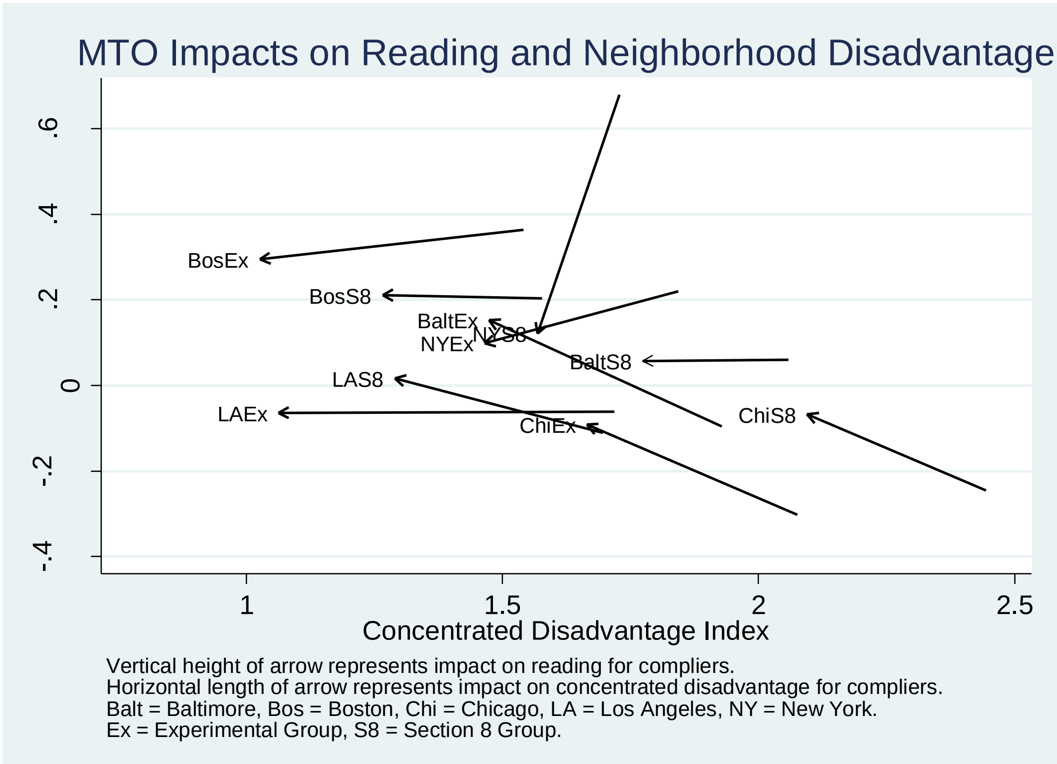
An alternative approach is suggested by Kling, Liebman and Katz (2007), who take advantage of the fact that there is variation across MTO cities and treatment groups in the degree to which MTO treatment group assignment affects different neighborhood attributes. They propose an instrumental variables (IV) design in which interactions between indicators for MTO city and MTO treatment group assignment are used as instruments for specific neighborhood attributes in a regression against whatever outcome measures are of interest. This design essentially asks whether those treatment groups in sites that experience the largest changes in particular neighborhood attributes are also the ones that experience the largest changes in outcomes. The power of this test is limited by the fact that we have just three randomized groups and five cities in MTO, so that the IV design essentially collapses the data into just 15 points. The test also assumes that the only reason that there are differences across cities in responses to the MTO treatment is because the amount of neighborhood change experienced by families in response to randomized MTO treatment group assignment varies across cities. With that caveat in mind, we find some suggestive evidence of non-linearity in the MTO data for reading scores, where the IV coefficient on the concentrated disadvantage index is .75 (standard error .47) and the coefficient on concentrated disadvantage index squared is -.26 (standard error .14). The coefficients and standard errors for math scores equal -.61 (.45) and .19 (.13). It is not clear exactly how much should be made of these results given that they are estimated fairly imprecisely.

While our analyses do not yield clear, convincing evidence for such non-linearities, it is important to note that our tests have relatively weak statistical power.

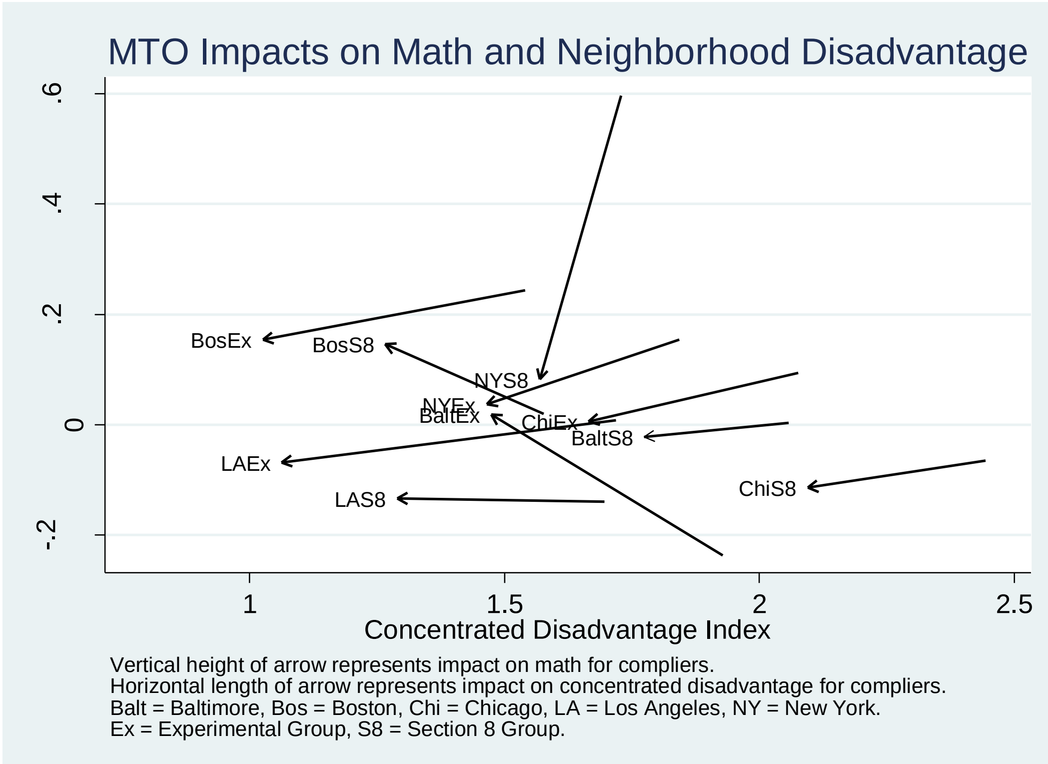
Appendix Figure 1
Propensity Adjusted Relationship Between Concentrated Disadvantage and Verbal Test Scores for African American and Hispanic Children in the PHDCN Study, Age Cohorts 6-12



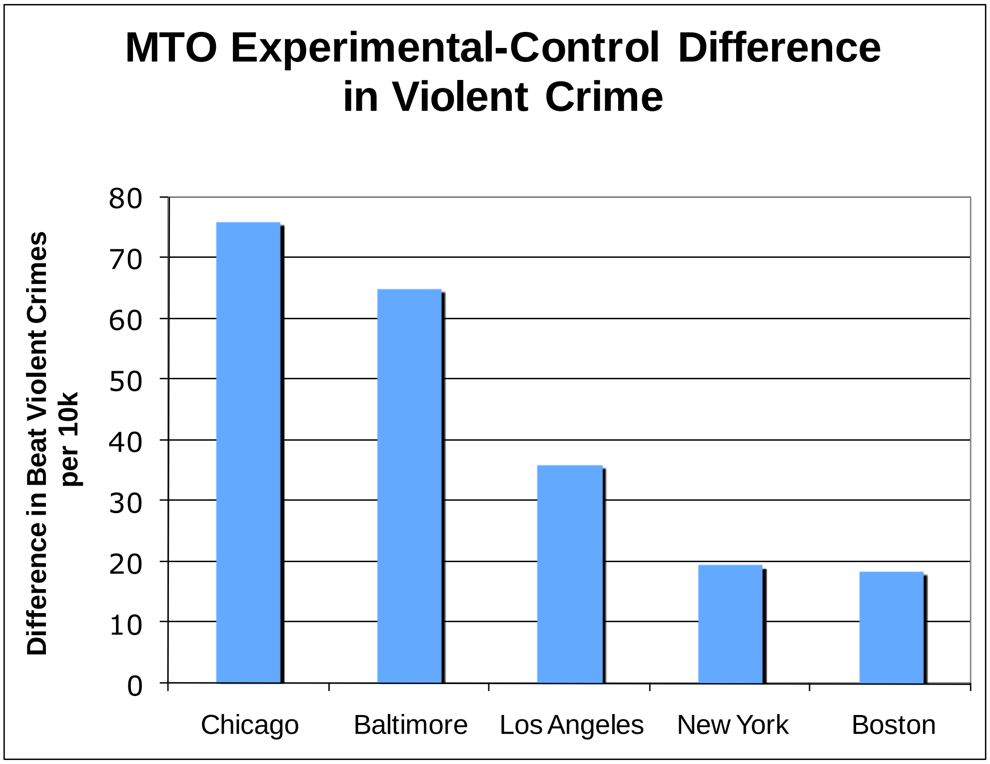
Appendix Figure 2: Relationship between concentrated neighborhood disadvantage and children's reading scores across MTO sites



Appendix Figure 3: Relationship between concentrated neighborhood disadvantage and children's reading scores across MTO sites



Appendix Figure 4:
MTO site-by-site results in effects of treatment assignment on beat-level local violent crime rates



Notes: The y-axis in the figure shows the difference in the average violent crime rate per 10,000 police beat or district residents for families assigned to the MTO experimental treatment group rather than the MTO control group (an intent to treat effect), for each of the MTO sites listed along the x-axis.

Appendix Table 1: Comparing Baseline Characteristics Across Study Samples

	Gautreaux	Public Housing Demolitions	PHDCN: African American	PHDCN: Hispanic	CHAC: Public Housing	CHAC: in MTO Tract at Baseline	MTO: Full Sample	MTO: Chicago Only	MTO: Chicago, Baltimore Only	MTO: NY, LA, Boston
Child Age	8.47	10.34	9.01	8.93	7.76	7.67				
		(4.01)	(2.52)	(2.49)	(2.21)	(2.25)				
Household Head Characteristics:										
Age	36.06		36.83	35.34	30.51	30.05	34.09	32.49	32.91	34.81
			(9.30)	(6.93)	(6.64)	(6.33)	(9.08)	(8.78)	(8.78)	(9.18)
African American	1.00	1.00	0.98	0.01	0.98	0.99	.67	.99	.99	.47
			(0.13)	(0.09)	(0.13)	(0.08)	(.4)	(.09)	(.12)	(.50)
Hispanic			0.00	0.95	0.01	0.00	.29	.01	.01	.46
			(0.04)	(0.22)	(0.09)	(0.03)	(.45)	(.08)	(.11)	(.50)
Employed			0.53	0.52	0.35	0.33	.27	.27	.26	.27
			(0.50)	(0.50)	(0.48)	(0.47)	(.43)	(.43)	(.43)	(.44)
Receiving Welfare	50.03		0.48	0.23	0.83	0.85	.74	.81	.81	.71
			(0.50)	(0.42)	(0.38)	(0.36)	(.43)	(.39)	(.39)	(.45)
Neighborhood Characteristics:										
Tract Poverty Rate		0.84	0.27	0.22	0.61	0.71	.50	.66	.58	.45
		(0.11)	(0.13)	(0.10)	(0.19)	(0.11)	(.14)	(.10)	(.15)	(.12)
Tract share black			0.76	0.13	0.89	0.99	.59	.99	.90	.39
			(0.29)	(0.18)	(0.24)	(0.06)	(.33)	(.04)	(.23)	(.21)
Tract share adults unemployed			0.17	0.11	0.33	0.39	.25	.40	.33	.20
			(0.07)	(0.04)	(0.13)	(0.11)	(.12)	(.08)	(.12)	(.08)
Tract share female-headed households			0.52	0.29	0.77	0.85	.65	.88	.80	.56
			(0.14)	(0.11)	(0.18)	(0.06)	(.20)	(.06)	(.18)	(.14)
Tract share persons on welfare			0.24	0.15	0.47	0.55	.23	.36	.26	.22
			(0.10)	(0.08)	(0.17)	(0.14)	(.13)	(.12)	(.16)	(.11)
Tract share persons under age 18			0.30	0.31	0.45	0.50	.38	.49	.40	.36
			(0.06)	(0.06)	(0.11)	(0.07)	(.11)	(.06)	(.13)	(.09)
Concentrated disadvantage index			2.20	0.70	3.00	3.39	2.18	3.16	2.74	1.84
			(1.11)	(0.85)	(0.77)	(0.33)	(.72)	(.29)	(.71)	(.46)
Concentrated disadvantage index (without % black)			1.93	0.84	2.25	2.56	1.69	2.34	1.99	1.51
			(1.18)	(0.87)	(0.61)	(0.31)	(.51)	(.27)	(.55)	(.38)

Notes: This table reports baseline household and neighborhood characteristics for the different studies that we review: Project on Human Development in Chicago Neighborhoods (Sampson, Sharkey and Raudenbush, 2008); Chicago public housing demolition study (Jacob 2004); Chicago CHAC voucher study for families living in public housing at baseline (Ludwig et al., 2010); and results from the Moving to Opportunity (MTO) study for different cities (Sanbonmatsu et al., 2006). The concentrated disadvantage index is a weighted average of several different census tract-level characteristics including tract share poor, tract share black, tract share unemployed, tract share households headed by a female, tract share on welfare, and share of the tract's population that is under age 18.

Appendix Table 2:
Control Means and Effects of Voucher-Induced Mobility at Follow-Up – Neighborhood Characteristics

	CHAC: Public Housing at Baseline	CHAC: in MTO Census Tract at Baseline	MTO: Full Sample	MTO: Chicago Only	MTO: Chicago and Baltimore Only	MTO: NY, LA and Boston
Tract Poverty Rate						
Control Mean	0.481	0.467	0.392	0.419	0.387	0.394
Voucher movers vs. non-movers	-0.274***	-0.336	-0.190*	-0.187*	-0.141*	-0.213*
	(0.094)	(0.259)	(.019)	(.069)	(.041)	(.018)
Tract share black						
Control Mean	0.837	0.912	0.548	0.857	0.848	0.371
Voucher movers vs. non-movers	0.028	-0.112	-0.023	0.04	-0.059	-0.01
	(0.091)	(0.287)	(.028)	(.089)	(.057)	(.029)
Tract share adults unemployed						
Control Mean	0.142	0.147	0.191	0.243	0.221	0.173
Voucher movers vs. non-movers	-0.0421*	-0.060	-0.074*	-0.071	-0.065*	-0.078*
	(0.025)	(0.053)	(.010)	(.039)	(.023)	(.009)
Tract share female-headed households						
Control Mean	0.327	0.353	0.558	0.649	0.64	0.51
Voucher movers vs. non-movers	-0.145**	-0.277	-0.133*	-0.077	-0.105*	-0.144*
	(0.060)	(0.183)	(.021)	(.072)	(.043)	(.020)
Tract share persons on welfare						
Control Mean	0.257	0.289	0.177	0.192	0.164	0.184
Voucher movers vs. non-movers	-0.113**	-0.239	-0.085*	-0.066*	-0.038	-0.109*
	(0.052)	(0.169)	(.011)	(.039)	(.022)	(.012)
Tract share persons under age 18						
Control Mean	0.402	0.412	0.357	0.387	0.359	0.355
Voucher movers vs. non-movers	-0.091**	-0.155	-0.054*	-0.096*	-0.055*	-0.051*
	(0.036)	(0.109)	(.009)	(.029)	(.017)	(.009)
Concentrated disadvantage index						
Control Mean	2.057	2.170	1.869	2.307	2.192	1.678
Voucher movers vs. non-movers	-0.548**	-1.012	-0.488*	-0.411	-0.399*	-0.530*
	(0.258)	(0.809)	(.067)	(.242)	(.145)	(.064)
Concentrated disadvantage index (without % black)						
Control Mean	1.357	1.408	1.409	1.59	1.482	1.366
Voucher movers vs. non-movers	-0.572***	-0.918	-0.466*	-0.444	-0.350*	-0.516*
	(0.215)	(0.648)	(.052)	(.19)	(.111)	(.051)

Notes: This table reports the effects of relocating using a housing voucher on different neighborhood characteristics reported at left; that is, each cell in the table represents the difference in average neighborhood characteristics for children who moved with a voucher versus the average for those children in the control group who would have moved had their families been assigned a voucher (the effect of treatment on the treated, or TOT). The voucher effect cells report the difference in average characteristics with the standard error underneath reported in parentheses. Each column reports results for a different study and/or sample within a study: Project on Human Development in Chicago Neighborhoods (Sampson, Sharkey and Raudenbush, 2008); Chicago public housing demolition study (Jacob 2004); Chicago CHAC voucher study for families living in public housing at baseline (Ludwig et al., 2010); and results from the Moving to Opportunity (MTO) study for different cities (Sanbonmatsu et al., 2006). The concentrated disadvantage index is a weighted average of several different census tract-level characteristics including tract share poor, tract share black, tract share unemployed, tract share households headed by a female, tract share on welfare, and share of the tract's population that is under age 18.

Appendix Table 3: City-wide crime rates for the 5 MTO cities

	1998 homicide rate	1998 UCR Part 1 violent crime rate	1998 UCR Part 1 property crime rate
Baltimore	47.1	2,420	8,527
Boston	6.1	1,327	4,924
Chicago	25.6	2,191	6,884
Los Angeles	11.8	1,359	3,714
New York	8.6	1,167	3,225

Table shows crime rates per 100,000 city residents. Source: U.S. Statistical Abstracts, 2000, Table 332. Violent crime rate for Chicago is taken from the Chicago PD annual report for 1998, rather than from the Statistical Abstracts, because the city's forcible rape figures are not reported in accordance with the national Uniform Crime Reporting guidelines and so not reported by the FBI (Chicago uses a broader definition than rape, to include all criminal sexual assaults, although only 2,387 out of the city's total 62,947 part 1 UCR violent crimes were criminal sexual assaults, so this should not have much impact on the comparability of the figures reported in the table). The population denominator we use for Chicago for 1998 is linearly interpolated from 1990 and 2000 decennial census figures.