NBER WORKING PAPER SERIES

EXPERIMENTAL ANALYSIS OF NEIGHBORHOOD EFFECTS

Jeffrey R. Kling
Jeffrey B. Liebman
Lawrence F. Katz

Working Paper 11577
http://www.nber.org/papers/w11577

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2005

This paper integrates material previously circulated in Kling and Liebman (2004), Kling et al. (2004), and Liebman et al. (2004). Support for this research was provided by grants from the U.S. Department of Housing and Urban Development, the National Institute of Child Health and Development (NICHD) and the National Institute of Mental Health (R01-HD40404 and R01-HD40444), the National Science Foundation (9876337 and 0091854), the Robert Wood Johnson Foundation, the Russell Sage Foundation, the Smith Richardson Foundation, the MacArthur Foundation, the W.T. Grant Foundation, and the Spencer Foundation. Additional support was provided by grants to Princeton University from the Robert Wood Johnson Foundation and from NICHD (5P30-HD32030 for the Office of Population Research), by the Princeton Industrial Relations Section, the Bendheim-Thomas Center for Research on Child Wellbeing, the Princeton Center for Health and Wellbeing, and the National Bureau of Economic Research. We are grateful to Todd Richardson and Mark Shroder of HUD; to Eric Beecroft, Judie Feins, Barbara Goodson, Robin Jacob, Stephen Kennedy, Larry Orr, and Rhiannon Patterson of Abt Associates; to our collaborators Jeanne Brooks-Gunn, Alessandra Del Conte Dickovick, Greg Duncan, Tama Leventhal, Jens Ludwig, Bruce Psaty, Lisa Sanbonmatsu, and Robert Whitaker; to our research assistants Ken Fortson, Jane Garrison, Erin Metcalf, and Josh Meltzer; and to numerous colleagues and three anonymous referees for valuable suggestions. Any findings or conclusions expressed are those of the authors. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Experimental Analysis of Neighborhood Effects
Jeffrey R. Kling, Jeffrey B. Liebman and Lawrence F. Katz
NBER Working Paper No. 11577
August 2005
JEL No. H43, I18, I38, J38

ABSTRACT

Families, primarily female-headed minority households with children, living in high-poverty public housing projects in five U.S. cities were offered housing vouchers by lottery in the Moving to Opportunity program. Four to seven years after random assignment, families offered vouchers lived in safer neighborhoods that had lower poverty rates than those of the control group not offered vouchers. We find no significant overall effects of this intervention on adult economic self-sufficiency or physical health. Mental health benefits of the voucher offers for adults and for female youth were substantial. Beneficial effects for female youth on education, risky behavior, and physical health were offset by adverse effects for male youth. For outcomes exhibiting significant treatment effects, we find, using variation in treatment intensity across voucher types and cities, that the relationship between neighborhood poverty rate and outcomes is approximately linear.

Jeffrey R. Kling
Firestone A-16-J
Princeton University
Princeton, NJ 08544-2098
and NBER
kling@princeton.edu

Lawrence F. Katz
Department of Economics
Harvard University
Cambridge, MA 02138
and NBER
lkatz@harvard.edu

Jeffrey B. Liebman
Kennedy School of Government
Harvard University
79 JFK Street
Cambridge, MA 02138
and NBER
jeffrey_liebman@harvard.edu

The residents of disadvantaged neighborhoods fare substantially worse on a wide range of socioeconomic and health outcomes than do those with more affluent neighbors. Economic models of residential sorting --partially motivated by these observed associations between neighborhood characteristics and individual outcomes -- suggest that inefficient equilibria can arise when individual outcomes are influenced by neighbors and individuals do not take their external effects on neighbors into account in their location decisions (e.g., Benabou, 1993).

It is hard to judge from theory alone whether the externalities from having neighbors of higher socioeconomic status are predominantly beneficial (from social connections, positive role models, reduced exposure to violence, and more community resources), inconsequential (only family influences, genetic endowments, and individual human capital investments matter), or adverse (from competition with advantaged peers and discrimination).[1] Empirical assessment of the importance of such externalities has also proven difficult using non-experimental data because individuals sort across neighborhoods for reasons that are likely to be correlated with the underlying determinants of their outcomes.

In this paper, we avoid the problem of endogenous neighborhood selection by using data from a randomized experiment in which some families living in high-poverty housing projects were offered housing vouchers to enable them to move to lower-poverty neighborhoods while others were not offered vouchers.[2] Thus our analysis provides direct evidence on the existence, direction, and magnitude of neighborhood effects for important socioeconomic and health outcomes in both adult and youth populations. The findings also bear on key housing policy decisions such as whether it is better for the government to provide housing subsidies tied to public housing projects or housing vouchers that can be used in the private-sector rental market.

[1] Jencks and Mayer (1990), Becker and Murphy (2000), Brock and Durlauf (2001)), Kawachi and Berkman (2003).
[2] Tenants in U.S. public housing and those using federal Section 8 housing vouchers both pay approximately 30 percent of their income in rent (Olsen 2003). The value of a voucher is the difference between 30 percent of income and the city's Fair Market Rent, set at the 40th percentile of area rents.

The research design used in this paper is based on comparisons of three groups to which households were randomly assigned in the Moving To Opportunity (MTO) social experiment, operated in five cities -- Baltimore, Boston, Chicago, Los Angeles, and New York -- by the U.S. Department of Housing and Urban Development. A control group received no new assistance, but continued to be eligible for public housing. A Section 8 group received a traditional Section 8 voucher, without geographic restriction. An experimental group received a Section 8 voucher, restricted for one year to a census tract with a poverty rate of less than 10 percent, and mobility counseling. Our sample consists of 4248 households assigned from 1994-97 at the five sites.

In 2002, extensive data were collected on outcomes from five key domains: economic self-sufficiency, mental health, physical health, risky behavior, and education. This paper provides the main results from MTO for adults and for youth ages 15-20 at all five sites an average of five years after random assignment, providing the most comprehensive experimental analysis to date of neighborhoods effects.[3]

1. Data and descriptive statistics.

The data for this study come from a baseline survey, from administrative data, and from an impact evaluation survey conducted in 2002 of one adult and up to two randomly selected children in each MTO household. The baseline survey was administered to household heads prior to random assignment. Administrative data on earnings and welfare benefits were obtained from state and county agencies in California, Illinois, Maryland, Massachusetts, and New York.[4]

[3] Additional results are available from Sanbonmatsu et al. (2004) who analyze reading and math test scores for children ages 6-20 and from Kling, Ludwig, and Katz (2005) who analyze arrest records for youth ages 15-25. The earlier single-site pilot studies of MTO are collected in Goering and Feins (2003).

[4] Four of the states provided individual-level earnings information on each MTO sample member who matched to the UI records. Massachusetts could provide the data only aggregated across groups consisting of at least 10 MTO individuals. Data was linked by Social Security Number (SSN). SSNs from each state of random assignment were linked to agency data only for that state. Earnings and welfare amounts were inflation adjusted to 2001 dollars using the Consumer Price Index for all urban consumers.

The 2002 survey had an effective response rate of 90 percent for adults and female youth and 86 percent for male youth.[5] All statistical estimates in this paper use weights.[6] The baseline covariates included in the regressions and supplemental details regarding outcome variables are given in the Appendix.

The Boston, Los Angeles, and New York families in our sample are mainly black or Hispanic; those in Baltimore and Chicago are nearly all black. Overall, 85 percent of the households are female-headed and either African-American or Hispanic. 98 percent of the sample adults were female, and 93 percent were ages 25-54 as of December 31, 2001. At the time of random assignment, one quarter of sample adults were employed, three quarters were receiving AFDC, more than half had never married, fewer than half had graduated from high school, and a quarter had been teenage parents. In a baseline survey, a majority said they wanted to move out of public housing "to get away from drugs and gangs."

Participants volunteered for this study, presumably because they were interested in moving out of their original high-poverty neighborhoods. Although this may be the most relevant population when considering incremental expansion of the use of housing vouchers to replace public housing, care should be taken in applying these results to populations with different characteristics. The experiment did not result in large clusters of moves to the same new

[5] An initial phase from January - June 2002 resulted in an 80% response rate for adults. At that point, we drew a 3-in-10 subsample of remaining cases and located 48% of them. The purpose of the subsampling was to concentrate our remaining resources on finding hard-to-locate families in a way that would minimize the potential for non-response bias. We calculate the effective response rate for adults as 80 + (1 - .8)*48 = 89.6. The effective response rate for youths is calculated in an analogous manner.

[6] The weights have three components (Orr et al., 2003). First, subsample members receive greater weight since, in addition to themselves, they represent individuals whom we did not attempt to contact during the subsampling phase. Second, youth from large families receive greater weight since we randomly sampled two children per household implying that youth from large families are representative of a larger fraction of the study population; this component does not apply to adults. Third, all individuals are weighted by the inverse of their probability of assignment to their experimental group to account for changes in the random assignment ratios over time. The ratio of individuals randomly assigned to treatment groups was changed during the course of the demonstration to minimize the minimum detectable effects after take-up of the vouchers turned out to be different than had been projected. This third component of the weights prevents time or cohort effects from confounding the results. Our weights imply that each random assignment period is weighted in proportion to the number of people randomly assigned in that period. Analyses of administrative data use only the third component of the weights.

neighborhoods; therefore it is unlikely to have had large effects on receiving neighborhoods or via "social multiplier" or "general equilibrium" effects (Manski 1993, Heckman 2001).

Families in the treatment groups had four to six months to find qualified housing and move using an MTO voucher. The fraction of treatment group families who used an MTO voucher to move -- which we refer to as the compliance rate -- was 47 percent for the experimental group and 60 percent for the Section 8 group. Compared to non-compliers (those in the treatment groups who do not use an MTO voucher), compliers (those who move using an MTO voucher) are younger and more likely to have had no teenage children at baseline, to have reported that their neighborhood was very unsafe at night, to have said that they were very dissatisfied with their apartment, to have been enrolled in school, and to have forecast that they would be "very likely" to find a new apartment if offered a voucher. Compliance rates differed substantially by site from a low of 32 percent in the Chicago experimental group to a high of 77 percent in the Los Angeles Section 8 group.[7]

To characterize the neighborhoods in which families lived and the differences in residential location for those who used an MTO voucher versus those who did not, Figure 1 shows several densities of neighborhood (census tract) poverty rates. The poverty rates are duration-weighted averages over locations lived at since random assignment, and use linear interpolation for poverty rates between the Census years of 1990 and 2000. Figure 1 indicates that experimental compliers lived in neighborhoods with significantly lower poverty rates than did controls, with nearly 60 percent living in neighborhoods below 20 percent poverty; Section 8 compliers also lived in lower-poverty neighborhoods, but their density is shifted by a more modest magnitude. The densities for experimental non-compliers, Section 8 non-compliers, and controls are quite

[7] The intensity of housing search assistance provided to the experimental group by the non-profits responsible for the counseling varied considerably across sites, as did the tightness of local housing markets.

similar to each other.[8]

Additional descriptive statistics of the residential locations are shown in Table 1. The experimental and Section 8 groups are both substantially less likely to live in very poor areas with visible drug activity, and somewhat more likely to live in areas with greater adult employment and a lower share of minority residents. Members of the treatment groups feel safer and are less likely to report a household member having been victimized by crime in the previous six months. The .82 average share minority for the experimental group tracts is indicative of the fact that, while families moved to lower poverty census tracts, these families did not move to distant white suburban areas. In the experimental group, only 16 percent moved 10 miles or more, and only 12 percent had an average tract share minority less than half.

2. Analysis.

We focus on fifteen primary outcomes for adults and fifteen primary outcomes for youth. Prior to examining the data, we decided to examine youth results pooled by gender and separately for females and males -- both because the prevalences for some outcomes differ greatly by gender and result in different statistical power to detect effects and because there had been some evidence of more beneficial effects for boys in earlier MTO research (Katz, Kling, and Liebman 2001; Ludwig, Duncan, and Hirshfield 2001) and in welfare reform research (Bos et al 1999). With fifteen outcomes, four population groups (adults, all youth, female youth, male youth) and two treatment groups, there are a total of 120 treatment effect estimates in this set.

To draw general conclusions about the experiment's results, we first present findings for summary indices that aggregate information over multiple treatment effect estimates (later we

[8] This implies that there was little selection of the type typically hypothesized, where compliers would have been more likely to have moved to lower-poverty neighborhoods even if they had not been offered a voucher (and the poverty distribution for controls would therefore exhibit greater density at lower neighborhood-poverty rates than would the density for noncompliers).

present estimates for specific outcomes). For example, we create an index of economic self-sufficiency that averages together five measures of employment, earnings, and public assistance receipt. The aggregation improves statistical power to detect effects that go in the same direction within a domain.[9] The index *Y* is an equally weighted average of z-scores of the components of the index, with signs of measures oriented as defined in the notes to Table 2 so that more beneficial outcomes have higher index scores. The z-scores are calculated by subtracting the control group mean and dividing by the control group standard deviation, so the value of the index has mean zero and standard deviation one by construction for the control group.[10]

We begin by estimating intent-to-treat (ITT) effects -- differences between treatment and control group means. Estimation of the ITT effect π_1 is from equation (1). Let *Z* be an indicator for treatment group assignment, and *X* be a matrix of baseline covariates.

(1) $Y = Z\pi_1 + X\beta_1 + \varepsilon_1$

X is included to improve estimation precision and to account for chance differences between groups in the distribution of pre-random assignment characteristics. Table 2 shows ITT results by domain and population group for indices -- with details on the measures included in each index provided in the notes to Table 2. The absolute magnitudes of the indices are in units akin to standardized test scores; the ITT estimate shows where the mean of the treatment group is in the distribution of the control group in terms of standard deviation units.

[9] O'Brien (1984) constructs a global test statistic for multiple endpoints with maximum power against the alternative that all effects have the same sign and effect size. An adaptation of the O'Brien approach is discussed in the Appendix and implemented in Kling and Liebman (2004) which uses seemingly unrelated regression effects for specific outcomes to estimate the covariance of the effects and calculates the mean effect size for groups of estimates in a second step. The average z-score index used in this paper is much simpler to work with, particularly for our results relating neighborhood poverty rates to outcomes. The two approaches yield identical treatment effects when there is no item nonresponse and no regression adjustment.

[10] If an individual has a valid response to at least one component measure of an index, then any missing values for other component measures are imputed at the random assignment group mean. This results in differences between treatment and control means of an index being the same as the average of treatment and control means of the components of that index (when the components are divided by their control group standard deviation and have no missing value imputation), so that the index can be interpreted as the average of results for separate measures scaled to standard deviation units.

For adults, the direction of effects is positive for both the experimental and Section 8 groups relative to the control group for all three domains: economic self-sufficiency, physical health, and mental health. The effect on mental health for the experimental group is much larger in magnitude than the others and is the only adult estimate that is statistically significant at the 5 percent level. For results pooling all youth, the direction of effects is positive for mental health and education, and negative for physical health and risky behavior. Again, the effects on mental health (for both the experimental group and Section 8 groups) are much larger and have p-values below .06. For the overall index averaging together all fifteen outcomes, the results in columns (i)-(iv) for adults and for all youth are positive in sign, but the magnitude is not large enough to reject a null hypothesis of no effect with 95 percent confidence. Thus, for adults and for all youth, the strongest evidence of effects from relocation to lower poverty neighborhoods is for the domain of mental health.

The overall results for youth average together estimates that substantially differ for female and male youth. Columns (v)-(viii) of Table 2 show large positive effects on mental health and risky behavior for female youth, and large negative effects on physical health and risky behavior for male youth.[11] This gender pattern in results was the opposite of what we expected.[12] Yet, as shown in columns (ix)-(x), the medium-term effects for females are more beneficial than for males for all four domains and in both treatment groups relative to the control group.

As a complement to the summary indices, we also examined results for each specific outcome that was a component of an index. Because the magnitudes of these separate outcomes

[11] These results are based on estimation of $Y = (1-G)(X\beta_{10} + Z\pi_{10}) + G(X\beta_{11} + Z\pi_{11}) + \upsilon$, where G is an indicator for gender, X includes household and individual-specific characteristics.

[12] All of the non-experimental papers that we are aware of showing gender differences in neighborhood effects find larger beneficial effects for boys from living in advantaged neighborhoods than for girls (examples include Entwisle, Alexander, and Olson , 1994; Ramirez-Valles, Zimmerman, and Juarez, 2002; and Crane, 1991). The predominant mechanism proposed in these studies is that boys spend more time hanging out in the neighborhood (rather than in the home), and therefore are influenced more heavily by the neighborhood.

are often easier to interpret than those of the summary indices, we show in Table 3 all outcomes with ITT effects significant at the five percent level. In addition to ITT effects, we also report the effect of treatment-on-treated (TOT) for these measures. We estimate this effect using the offer of an MTO voucher as an instrumental variable for MTO voucher use, so Z is the excluded instrument for an indicator *D* of compliance in two stage least squares estimation of (2).[13]

(2) $Y = D\gamma_2 + X\beta_2 + \varepsilon_2$

The TOT parameter γ_2 is equal to the ITT parameter divided by the regression-adjusted compliance rate. This TOT approach relies on the assumption that there was no average effect of being offered an MTO voucher on those who did not use an MTO voucher, which we believe is a reasonable approximation, but not strictly true.[14] Under this assumption, we can assess the average magnitude of the effect of the voucher offer for those who complied and used an MTO voucher to move to a lower-poverty neighborhood.

As an example, results for the specific outcome of adult obesity, using a standard body mass index cutpoint (BMI ≥ 30 kg/m^2), are shown in the first row of Table 3. The results in column (i) compare the experimental and control groups (E-C), and the ITT effect on obesity is a five percentage point reduction in obesity for the experimental group relative to the control group.

[13] We interpret the 2SLS results as treatment-on-treated estimates rather than local average treatment effect (LATE) estimates (Imbens and Angrist 1994) because the endogenous variable is use of a voucher offered by the MTO program, and MTO vouchers are never offered to the control group; there are no always-takers in the terminology of Angrist, Imbens, and Rubin (1996). Over time, some control group members do receive housing vouchers from other sources, but they tend to receive them significantly later than treatment group members do. Therefore, in our TOT estimates, we do not define control group voucher recipients as having been treated. We do interpret both our ITT and TOT estimates as averages of heterogeneous treatment effects across individuals.

[14] For the experimental group, this assumption implies that the later outcomes of households who met with a housing mobility counselor were not affected by the counselor if that household did not make a subsidized move through the MTO program. For both treatment groups, this assumption implies that the experience of housing search induced by assignment to a treatment group did not affect later outcomes if that household did not make a subsidized program move. For noncompliers, we believe that the effects of mobility counselors (who mainly provided housing advice and not general social services) on self-sufficiency and health outcomes are likely to be orders of magnitude smaller than the effects of moving to a new residential location. The TOT approach also requires that the control group was not affected by the experience of losing the voucher lottery, something we view as a reasonable approximation.

Assuming no effect on those who did not use an MTO voucher to move, the TOT effect was ten percentage points.

Since outcomes are directly observed for treatment group compliers and we have a TOT estimate, we can estimate the mean level of each outcome for those in the control group who would have complied if they had been offered a voucher -- which we refer to as the control complier mean or CCM (Katz, Kling, and Liebman 2001) -- based on relation that CCM = E[Y | Z=1, D=1] - γ.[15] For the fraction obese, the CCM in column (iv) is estimated to be .502.[16] Thus a ten percentage point change would be a decline in relative risk among compliers of 21 percent and a relative odds ratio of .66 for obesity for experimental relative to control compliers.[17]

These effects on adult obesity are of substantial magnitude, yet they are the smallest in relative risk and relative odds among the 13 binary outcomes in Table 3. The two continuous outcomes in Table 3 (the z-scores from the K6 distress index) have TOT effect sizes of .2 and .5 standard deviations. Thus, each of the TOT effects in Table 3 appears to be of a substantively important magnitude.

Another metric in which the magnitude of the results can be assessed is in terms of the association between a change in the neighborhood poverty rate (W) and the change in an outcome.[18] Thus in equation (3) the key parameter is γ_3, the OLS coefficient on poverty rate.[19]

---

[15] For binary outcomes, sampling variation can produce negative estimates of the CCM. Our analytic method assumes that the there would have been the same fraction of noncompliers in the control group as were observed in the treatment group, and that these noncompliers had exactly the same outcome prevalence as treatment noncompliers. In any one sample this method may produce a negative estimate of a CCM if a particular realization of the treatment noncomplier mean is higher than the realization of the control noncomplier mean, even though the method is unbiased in repeated sampling. In our results, we report a CCM of zero when the CCM estimate for a binary outcome is negative.

[16] BMI = 30 for a woman five feet four inches and 175 pounds. Nationally for ages 18-44, 33 percent of black women and 22 percent of Hispanic women have BMI ≥ 30 (Lucas, Schiller, and Benson 2004).

[17] The odds of obesity for control compliers are .502/(1-.502) = 1.01. The odds for experimental compliers are .399/(1-.399) = .664. The relative odds are .664 / 1.01 = .659.

[18] This approach permits direct comparison with the large nonexperimental literature. In addition, sociological threshold models (Granovetter, 1978) and economic models of sorting across neighborhoods, schools, and

(3) $Y = W\gamma_3 + X\beta_3 + \varepsilon_3$

OLS estimation of this regression may be biased by endogenous residential location as previously discussed, but treatment group assignment can be used to form instrumental variables. Consider a regression using all sample members, regardless of MTO group. With a set of excluded instruments containing site-by-treatment interactions and when *X* contains only site indicators, then the 2SLS estimate of γ_3 is the line fit through a scatterplot of 15 outcome and poverty rate means for the three random assignment groups in each of five sites, normalized so that each site has mean zero. This is depicted graphically in Figure 2, with four panels for four summary indices (adult mental health, female youth mental health, female youth overall, and male youth overall). The figures show that there is a consistent pattern across the sites and groups that larger differences in poverty rates (relative to the site mean) are associated with differences of larger magnitude in outcomes. The relationship between poverty rate and outcomes appears fairly linear, and in each case, a test of the overidentifying restrictions (Davidson and MacKinnon, 1993) has a p-value of less than .30 indicating that the data are consistent with a linear model.[20]

Estimates based on equation (3) are given in Table 4 for selected outcomes. The first column shows OLS estimates using data for the control group only – results illustrative of the approach taken in the non-experimental literature on neighborhood effects (e.g., Brooks-Gunn et. al., 1997) that could have been applied in analysis of this population without the random assignment

---

classrooms (Arnott and Rowse, 1987, de Bartolome, 1990; Fernandez and Rogerson, 1996; Benabou, 1993) hinge on the exact form of the relationship between neighborhood or peer groups characteristics and individual outcomes.

[19] We interpret the coefficient on neighborhood poverty rate as the effect of moving a neighborhood with a lower poverty rate and other associated differences in neighborhood characteristics, and not as the effect of changing the poverty rate while holding other characteristics of the neighborhood constant.

[20] We have also used another diagnostic technique for examining the form of the relationship between the outcomes and poverty rates in each group-site – examining augmented partial residual plots, recommended by Mallows (1986) to detect nonlinearity in the relationship of W and Y. This analysis did not suggest any evidence of nonlinearities in our estimated relationships.

of vouchers. The second column shows 2SLS estimates of γ_3 for the entire sample with site-group interactions as excluded instruments for the neighborhood poverty rate *W*, and a full set of covariates in *X*. The 2SLS estimates bear little relation to the OLS estimates, implying that endogeneity is a substantial issue for non-experimental approaches to these data. Differences in neighborhood poverty rates of ten percentage points (roughly the treatment-control difference in poverty rates in Table 1) are associated with outcome effect sizes similar in magnitude to those for the ITT effects in Table 2.

To test the hypotheses that differences in poverty rates had the primary effects on outcomes as opposed to simply using an MTO voucher to move out of public housing, we also enriched *W* in equation (3) to include both the poverty rate and an indicator for compliance (*D*), with results reported in columns (iii) and (iv) of Table 4. Comparing columns (ii) and (iii), results without and with controls for compliance are quite similar for female youth and are within sampling error for adult mental health and for male youth (accounting for the covariance of the estimates). For the more precisely estimated models (adult mental health, female youth overall, male youth overall), the coefficients on poverty rates are large both in absolute magnitude and relative to their standard errors, while the coefficients on compliance are wrong-signed and small relative to their standard errors -- providing some evidence that the poverty rate effect was more important than the "move per se" effect. We have also examined other models with two endogenous variables, such as poverty and poverty-squared, intercept shifts in poverty rates, and kink points in poverty rates; while there is no evidence of nonlinearities in these models, the research design has little power to identify these effects.

3. Discussion

Adult economic self-sufficiency. The idea that residence in a distressed community can limit an individual's economic prospects has been powerfully advanced by Wilson (1987), and the related hypothesis that proximity to employment is important has its roots in the spatial mismatch hypothesis of Kain (1968). However, we found no significant evidence of treatment effects on earnings, welfare participation, or amount of government assistance after an average of five years since random assignment. As shown in Figure 3, the fraction of MTO adults with positive quarterly UI earnings increased from less than 25 percent in early 1995 to more than 50 percent in 2001. The time-patterns are, however, similar for the three randomly-assigned groups.[21]

In analyses shown in the Appendix, we explore possible reasons for the lack of effects. There do not appear to be important differences across the three MTO groups in job accessibility, to the extent that aggregate employment growth in establishments at the zip code level reflects available job vacancies.[22] Also, the MTO intervention had only small impacts on job-related social networks. While we found that the intervention modestly increased the fraction of sample members who "had a friend who graduated from college or earned more than $30,000 a year," we also found that only about eight percent of the sample "found a job through someone living in

[21] The strength of the U.S. labor market, welfare reform, and the declining fraction of sample members with preschool-age children at home are the most likely explanations for this upward trend. In analyses shown in the Appendix we found similar results based on survey and administrative data, suggesting little bias in self-reports. Given that the name of the demonstration is "Moving To Opportunity" and it was promoted by the government as a pathway to better jobs, one might have expected employment and earnings to be the most likely outcomes to be exaggerated by treatment group members if they were trying to "tell us what they thought we wanted to hear."

[22] It is somewhat surprising that the MTO intervention – which assisted families in moving out of some of the most concentrated pockets of poverty in the U.S. – had no discernable overall effects on employment, given that a recent comprehensive survey (Ihlandfeldt and Sjoquist, 1998) concludes that the empirical evidence overwhelmingly supports the spatial mismatch hypothesis that inner-city low-skilled minority workers have relatively weak access to jobs because job opportunities are disproportionately in suburban areas and housing market discrimination plus commuting costs create barriers that prevent minorities from reaching suburban jobs. It turns out, however, that the MTO experiment provides only a weak test of this aspect of spatial mismatch since the effects on distance moved and on local area job growth are small. If spatial mismatch is construed more broadly to encompass residence in distressed, unsafe, crime-ridden communities inhibiting access to jobs outside these communities, the results do indicate that moving to communities that are substantially safer does not have detectable effects on employment.

their neighborhood such as a friend, relative, or acquaintance" -- and that this proportion did not vary across MTO groups.

We also explored whether effects differed by baseline characteristics, and in general results do not differ appreciably by these characteristics.[23] In results shown in the Appendix, we find some suggestive evidence using UI data of interesting dynamics in the treatment effects on employment and earnings for younger adults, with initial negative treatment effects in the first two years after random assignment fading away over time for the Section 8 group and turning positive and substantial in the fourth and fifth years after random assignment for the experimental group.[24]

Adult physical health. Our early work at the Boston site (Katz, Kling, Liebman, 2001) suggested that the MTO intervention may have had important health impacts and led us to expand the outcomes studied beyond the economic and housing outcomes that were the original focus of the MTO demonstration. In the much more extensive health data gathered in the current study, there was not a broad pattern of physical health improvements for the treatment groups. The intervention did not have statistically significant effects on self-reported overall health, hypertension, asthma, or trouble carrying groceries or climbing stairs and there was not a statistically significant effect on the physical health index in Table 2.[25]

There was a large and statistically significant effect on obesity (see Table 3), possibly related to the reduced psychological distress and increase in exercise and nutrition that we also observed

[23] Results are given in the Appendix by age. Results by education level, gender composition of the household, and whether the household head was a prime-age minority single mother are available from the authors.

[24] Examining effects by age was an exploratory and not a confirmatory exercise, and we note that this type of searching for significant effects in subgroups raises the chance of concluding that there are statistically significant results even when the null hypothesis of no effects is true.

[25] Further exploration of the data by age showed a positive and significant impact of the MTO experimental treatment on our summary measure of physical health outcomes for the younger adults and no significant overall impact for older adults. These health impacts come from aggregating five consistently-signed estimates with small magnitudes rather than from a large effect on any one measure. This result, along with the suggestive evidence of employment gains among younger adults, leads us to speculate that the habits and behaviors of younger adults may be more malleable and therefore more responsive to a change in residential environment.

for the treatment groups. The interpretation of the obesity results depends largely on one's reason for focusing on obesity. If a researcher was searching for social experiments about the effects of neighborhoods on obesity (and this is the only one), focusing on the t-statistic of 2.2 and the associated per-comparison p-value of .03 is highly suggestive of an important effect of residential location on obesity. If a researcher is searching through the MTO results for significant effects, however, we suggest some caution in interpreting the obesity results. When using a method that focuses on particular results from among many outcomes of a family because of their t-statistics, there is considerable likelihood that the obesity results could have been observed by chance under a joint null hypothesis of no effects for all estimates in that family.[26]

Adult mental health. In contrast to the results for physical health, the adult mental health results were quite consistent across specific measures (distress, depression, anxiety, calmness, sleep) in finding beneficial effects for the experimental group relative to the control group. This consistency led to the large mean (ITT) effect size estimate of .08 standard deviations for the adult mental health summary measure in Table 2. The confidence level that the results are not due to chance is quite high under a method where the focus on mental health is determined exogenously (leading to per-comparison inference) or endogenously from the high t-statistic (leading to familywise inference).[27] The magnitude of the mental health results – for example a 45 percent reduction in relative risk among compliers of scoring above the K6 screening cutpoint

[26] The probability that the second largest t-statistic among 30 adult estimates is 2.2 or higher under the joint null hypothesis of no effect is a familywise adjusted p-value of .80, which is indicative of the fact that there is little power to reject the joint null hypothesis for specific outcomes. The adjusted p-values throughout this paper are based on a bootstrap procedure accounting for covariance among estimates, using a method adapted from Westfall and Young (1993) as described in the Appendix.

[27] We focus our familywise inference on the summary indices, which reduces the dimension of the inference problem. For adults, we focus on 8 ITT estimates for summary indices in Table 2. The familywise adjusted p-value for the mental health index t-statistic being 2.8 or higher (i.e., the largest t-statistic on estimates for 8 adult indices) under the joint null hypothesis of no effect on any of the 8 adult indices in Table 2 is .06.

for serious mental illness (Kessler et al., 2003) -- is comparable to that found in some of the most effective clinical and pharmacologic mental health interventions.[28]

The overall pattern of adult results -- with the agreement of estimated effects based on self-reports and administrative records of economic outcomes, with effects concentrated in the single domain of mental health, and with mental health effect sizes systematically related to changes in neighborhood poverty rates in Figure 2A and in Table 4 -- indicates that there are beneficial impacts on mental health of moving to less distressed neighborhoods. In addition, this pattern is contrary to a model in which the unrestricted choice of the Section 8 group should have led to better outcomes than the restricted choice of the experimental group. Based in part on evidence from the extensive qualitative interviews that have been done with MTO participants and the strong associations shown in the MTO quantitative research, we believe that the leading hypothesis for the mechanism producing the mental health improvements involves the reduction in stress that occurred when families moved away from dangerous neighborhoods in which the fear of random violence influenced all aspects of their lives (Kling, Liebman and Katz, 2005; Popkin, Harris and Cunningham, 2002).

Female youth outcomes. Teenage youth are often seen as the population most at risk from the adverse effects of high-poverty neighborhoods. In this study, 15 specific outcomes were assessed for youth within the four domains of physical health, mental health, risky behavior, and education. A summary index of all 15 outcomes shows large benefits for female youth in the

[28] In a study often cited as an exemplar of an effective clinical intervention, Wells et al. (2000) analyzed outcomes of depressive patients randomized to obtain usual care or improved quality care (better training of medical staff and better follow-up with patients). 12 months later, the fraction with depressive symptoms in the quality improvement group was .42, while the fraction was .51 in the usual care group – a reduction in relative risk of 18 percent. A meta-analysis of clinical trials of medications for major depressive disorder found that on average 50 percent of patients receiving an active medication showed improvement compared with 29 percent receiving a placebo -- a reduction in relative risk of 30 percent (Walsh et al, 2002). The reduction in relative risk for major depressive episode among MTO experimental compliers was also 30 percent (Appendix Table F5). Improved mental health for the treatment groups relative to the control group was a mechanism that we had hypothesized might increase employment and earnings. Although the effects on mental health reported here are large, Kling et al. (2004) calculate that these are unlikely to translate into effects on earnings large enough to detect.

experimental and Section 8 groups relative to the control group. This pattern of beneficial effects is quite consistent across outcomes, with 13 of 15 outcomes having the sign of a treatment effect in a beneficial direction for both treatments. The magnitudes of the effects are largest for mental health, still substantial for education and risky behavior, and small for physical health. For example, the experimental compliers have a relative risk of serious generalized anxiety symptoms 70 percent lower than the control complier mean.

In terms of assessing statistical significance, we adopt a similar framework as with the adult outcomes. If there is ex-ante interest by a researcher in a particular estimate, then the per-comparison p-value for that estimate is appropriate. When considering the many estimates for youth simultaneously, there is a high probability of observing a few large estimates due to sampling variability even if there were no true effect. To account for these multiple comparisons while restricting the set to a manageable size, we considered inference for three youth subgroups (all, female, male), two treatments (experimental, Section 8), and five domains (physical health, mental health, risky behavior, education, and overall) -- which correspond to the 30 estimates in columns (iii)-(viii) of Table 2. We calculated familywise adjusted p-values, similar to Bonferroni corrections, but adjusted for the ordering of the tests and the covariance of the estimates as described in the Appendix. The estimate in this set with the largest t-statistic was the overall summary index for the experimental group, and the probability of observing an effect this large or larger as the maximum of the 30 estimates was less than .001 under the joint null hypothesis of no effects for the 30 estimates. The familywise adjusted p-value was .003 for the experimental group mental health index. Based on these calculations, we conclude that the overall pattern and the mental health result for the experimental group were quite unlikely to have occurred by chance even if one focused on these results because of their large t-statistics.

Male youth outcomes. The results for the overall summary index of male youth outcomes are of almost exactly the same magnitude as for female youth but with the opposite sign, implying more adverse effects in the treatment groups than in the control group. Among specific outcomes, the effects are largest for injuries and for substance use, as shown in Table 3, leading to large effects for the male physical health and risky behavior summary indices in Table 2. These summary index measures for males have highly significant per-comparison p-values, although familywise adjusted p-values were greater than .05.

There are a number of issues that complicate interpretation of the results for male youth. First, males in the treatment groups exhibited more behavior and other problems at baseline than did those in the control group (there were no such baseline differences among females).[29] This imbalance appears largely due to the random sampling that occurred when we subsampled children for our interviews, rather than to survey attrition or imbalance in the original random assignment.[30] The key question for our analysis is whether our regression controls for baseline covariates are sufficient to adjust for the imbalance. In many situations where imbalance is due to attrition or some other systematic reason, it is intuitive that unobservable variables are also imbalanced; therefore controlling for observable differences might not be sufficient. In our case, however, the source of imbalance is largely randomness. Therefore there is less reason to be

[29] This result comes from a summary index of baseline covariates constructed in the same manner as the outcomes indices (normalizing the control group for each gender to mean zero and standard deviation one). The index includes variables (collected prior to random assignment) for age, gifted classes, school suspension, problems at school, behavior problems, learning problems, physical activity problems, and other medical problems. To orient the index, the signs are reversed for all items except gifted classes.

[30] For the following E-C and S-C refer to ITT differences in the baseline covariate index in standard deviation units for male youth (with standard errors). For all 1604 male youth ages 15-20 in MTO households, the E-C difference was -.022 (.031), and the S-C difference was -.019 (.034). For the 923 male youth we attempted to survey (drawing two children per household and a 3-in-10 subsample of initial nonrespondents), the E-C difference was -.133 (.053), and the S-C difference was -.108 (.055). For the 879 youth with whom we completed surveys, the E-C difference was -.158 (.054), and the S-C difference was -.129 (.056). Since there was less than half a standard error difference, respectively, between the estimates for all male youth for whom surveys were attempted and for those completed but large imbalance between treatment groups in baseline covariates for those attempted, we conclude that the imbalance was largely driven by random sampling. For comparison with the 928 female youth surveyed, the E-C difference was -.021 (.044), and the S-C difference was -.047 (.055).

concerned about imbalance in unobservables, and we suspect that our regression adjustment is sufficient to remove most of the potential bias.[31] Moreover, in analysis of administrative arrest data using the full set of MTO youth (with little imbalance in covariates and no survey attrition), Kling, Ludwig, and Katz (2005) find adverse treatment effects for males and beneficial effects for females, supporting out conclusion that gender differences in effects are substantial and not simply an artifact of sampling issues.

A second issue is that the rates of some adverse outcomes in the control group seem implausibly low. For example, the proportion of non-sports injuries for male youth in the control group is barely half as high as the injury rate for females, whereas our analysis of National Health Interview Survey data found non-sports injury rates for male youth over 30 percent higher than for female youth. Moreover, assuming that the injury rate among control noncompliers is the same as for treatment noncompliers implies a control complier mean of less than zero. Both of these facts are consistent with a low realization of injury rates for the particular sample and time period for which we have data; we speculate that the injury rate for males would be at least as high for males as for females if we were to run the experiment again. The rates of substance use for males in the control group are also low relative to demographically similar individuals in national data, whereas there are smaller differences between MTO controls and national data for females.[32] Our interpretation of the results is that issues such as random covariate imbalance,

[31] It turns out that the estimates for injuries are scarcely different with and without regression adjustment. In contrast, risky behavior estimates are affected by regression adjustment. ITT estimates without regression adjustment for the four risky behavior outcomes are 0.3 to 1.7 percentage points larger for the experimental group and 1.8 to 4.9 percentage points larger for the Section 8 group than the regression-adjusted estimates.

[32] In the Appendix, we describe our method for producing results for the National Longitudinal Survey of Youth 1997 (NLSY97) that are adjusted to match the demographics of the MTO sample. The following results in parentheses are for the MTO control group, the adjusted NLSY97, and the unadjusted NLSY97 respectively. Marijuana (females: .13, .13, .16; males: .12, .21, .18). Cigarettes (females .20, .26, .33; males: .13, .29, .33). Alcohol (females: .23, .28, .43; males: .13, .30, .45). The deviations of the MTO controls from the adjusted NLSY97 across these three outcomes (in parentheses, respectively) are substantially smaller for females (0, .06, .05) than for males (.09, .16, .17), which is consistent with a random draw of unusually low substance use among males in the control group.

survey attrition, and surprisingly low prevalence of adverse outcomes among controls may exaggerate the magnitude of the adverse effects for males, but that these factors are not large enough to account for the differences in effects between males and females.

Understanding gender differences. We collected extensive data on mediating factors that could potentially help in determining the mechanisms responsible for the observed results.[33] Given the gender differences in our youth results, an important question is whether there were gender differences in any mediating factors. In brief, we find that there were large effects of the treatment on neighborhood characteristics and smaller effects on school characteristics, but no significant differences by gender in the treatment effects on neighborhoods or schools (see the Appendix for details). We conclude that female youth and male youth in the treatment groups responded to similar new neighborhood environments in different ways.

Several mechanisms could potentially explain the gender differences in outcomes. There has been a broad trend over the past two decades of gains in education and employment for minority women – gains that have not been shared by minority men (Altonji and Blank 1999). Moves through MTO may remove barriers to benefiting from these gains for female youth, whereas the male youth have poor prospects even in lower poverty neighborhoods. It is also likely that girls suffer disproportionately from domestic violence and sexual abuse (Popkin, Harris, and Cunningham, 2002), and the MTO intervention may have reduced their exposure to such events – providing benefits from the moves for girls that were not nearly as relevant for boys. We do find some statistically significant evidence that female youth are more likely to have three or more adult role models to whom they are comfortable talking about their problems, and that the mean effect size on a summary index of adult contact measures was significantly higher for

[33] Because the MTO intervention changed so many aspects of neighborhoods and housing environments simultaneously, the research design was not ideal for determining the mechanisms that may have produced the effects on outcomes that we observed. Interventions that change on feature of the environment at a time would more directly evaluation particular mechanisms.

female youth than for male youth. While the probability is quite high of observing at least one effect this large under a joint null hypothesis of no gender differences in the many mediating mechanisms we examined, our interpretation is that differences in adult contact are the most likely contributor to at least part of the gender difference in effects from among the mechanisms about which we have data to examine.[34]

For males, a two-audience signaling process (Austen-Smith and Fryer 2005) could encourage them to avoid peer sanction through participation in deviant group activities rather than engage in more pro-social behaviors that are ultimately valued by employers-- a process which could be more important when there is more uncertainty about social group affiliation due to greater racial, ethnic, or economic diversity among peers (as there would tend to be for youth in families using MTO vouchers). Fryer and Torelli (2005) find a stronger negative association between popularity and grades among high achievers for black males than for black females and that this difference is larger in less segregated schools; although this gender difference was concentrated among those with high grades and was not found in full population, their result is consistent with greater peer sanction against the pro-social activities of black males than those for females for some subgroups. For MTO, we have evidence on the prevalence of visits to baseline neighborhoods (that may have been negative influences through signaling to a group or through lack of parental supervision); the sign is consistent with males making more visits, but the male-female difference is insignificant. Another possibility is that relative deprivation or being at a lower point in the ability distribution in the new schools than in the pre-move schools has particularly adverse effects for males.[35]

[34] We also found that the beneficial effects on adult mental health outcomes overall (such as distress and calmness) were not evident for adults with male youth in their households, with the difference for male-female youth being significant. This finding could be the effect of adverse male outcomes rather than the cause.

[35] Also, male youth may have less effective coping strategies in stressful situations (Zaslow and Hayes, 1986; Coleman and Hendry, 1999; Kraemer 2000), and the disruption of moving itself may have been greater for male

Reconciling OLS and 2SLS estimates. If low neighborhood poverty rates were beneficial and the neighborhood selection process operated such that people with unobserved characteristics associated with good outcomes tended to locate in lower-poverty neighborhoods, these assumptions would predict that the OLS estimates of effects of higher poverty rates would be more adverse than 2SLS estimates. In Table 4, however, we found OLS estimates (based on the control group only) were smaller in absolute magnitude than 2SLS and often of opposite sign. The implied selection process is that adults and families with female teenagers likely to have adverse outcomes tended to move to low poverty neighborhoods, and families with male teenagers likely to have beneficial outcomes tended to move to low-poverty neighborhoods.[36] The complexity of these selection patterns implies that it will not in general be possible to identify the direction of bias in non-experimental studies of neighborhood effects.[37]

Younger Children. Sanbonmatsu et al. (2004) hypothesized finding greater effects of the MTO treatment on children who are younger than the youth whose outcomes we study in this

---

youth. However, at least two strands of evidence run counter to this hypothesis. First, mobility rates were slightly higher among households with female youth than those with male youth. Second, the adverse effects for males did not manifest themselves right after the initial moves, as predicted by a simple mobility disruption model, but only after several years. In studies of single MTO sites 1-3 years after random assignment, there were either no gender differences in effects reported, or more beneficial effects for males than females (Katz, Kling, and Liebman, 2001; Leventhal and Brooks-Gunn, 2003a, 2003b). In administrative arrest data, the adverse effects on male property crime are not found in the first two years after random assignment, but in the third and fourth years (Kling, Ludwig, and Katz, 2005). It is also possible that the destination neighborhoods lacked some of the institutional support for at-risk boys that might have been present in the origin neighborhoods and that male youth lacked male role models of their race or ethnicity in the new neighborhoods (given that nearly all MTO families are headed by single mothers, a female MTO youth will generally have an adult same-sex role model). We find however that the MTO intervention had no significant impacts on parenting practices, peer characteristics, school engagement, or access to health care for either boys or girls.

[36] These implications about the pattern of residential sorting are borne out within the treatment groups as well. For nearly all outcomes, the compliers are more similar to the noncompliers than to the control group. For female youth and adults, this pattern can only be consistent with beneficial treatment effects if compliers were people who otherwise would have had poor outcomes. And for male youth, this pattern can only be consistent with adverse treatment effects if compliers were people who would otherwise have had good outcomes. We examined the predictors of compliance and of average tract poverty rate separately for each group. Consistent predictors of greater compliance include younger adult age, smaller household size, and dissatisfaction with original neighborhood. No characteristics except site were consistent predictors of poverty rates across all three of the groups. Unfortunately, none of these predictors shed much light on the selection pattern.

[37] The standard assumption of positive selection may be more applicable to populations where the frequency is lower of recent moves and attendant transitory shocks or where there is greater scope for differences in unobserved characteristics than in this MTO sample of largely female-headed minority households from public housing.

paper because the younger children (such as those ages 6-10) will have had less lifetime exposure to the high poverty neighborhoods, but they found no statistically significant treatment effects for younger children on reading test scores, math test scores, or behavior problems. It is also the case, however, that the main outcomes for which we found large treatment effects for teenage youth – mental health problems and risky behavior – have very low prevalence at younger ages, and it is therefore too early to tell whether the outcomes of the younger children will be different.

4. Conclusion

Using a housing voucher lottery which caused otherwise similar groups of families to reside in very different neighborhoods, we have investigated the effects of moving out of some of the highest poverty neighborhoods in the United States on outcomes for adults and teenage youth. Our findings – no effects on adult economic self-sufficiency, improvements in adult mental health, beneficial outcomes for teenage girls, adverse outcomes for teenage boys – have three important implications.

First, housing mobility by itself does not appear to be an effective anti-poverty strategy – at least over a five-year horizon. The MTO demonstration program was motivated by theories and non-experimental empirical results suggesting that there would be large economic gains from moving to lower-poverty neighborhoods. However, we found no evidence of treatment effects on adult earnings or welfare participation. Whether economic gains begin to appear in the longer run, particularly among MTO children, remains to be seen.

Second, even in the absence of economic gains, policies that move families out of distressed public housing projects using rental vouchers are likely to have benefits that significantly exceed their costs. Because the MTO intervention produced large mental health improvements and

because other research suggests that it is cheaper to provide a unit of subsidized housing with vouchers than in a public housing project (Olsen 2000), an offer of a housing voucher is likely to pass the Kaldor-Hicks criterion --the gains to those who benefit would be large enough to hypothetically compensate those who experience adverse effects and still be better off. We note, however, that spillovers onto neighborhoods to which these families moved remain unknown. If there were large negative spillovers this conclusion could be reversed. In addition, the largely offsetting male and female youth results complicate the welfare analysis.

Third, substantively important neighborhood effects do exist, but only for some outcomes. Teenagers – the population often thought to be most affected by neighborhood conditions – exhibited effects on the broadest range of outcomes. The evidence that effects of housing vouchers appear to accrue from changes in neighborhood characteristics rather than from moves per se suggests that interventions which substantially improve distressed neighborhoods could have effects as least as large as those observed from moving to lower-poverty neighborhoods.[38] Although numerous non-experimental studies document strong associations between neighborhood characteristics and individual outcomes, these associations appear to be much weaker in the studies with the most credible identification strategies.[39] Because the current study used randomization to solve the selection problem, because it studied families who made very large moves as measured by changes in neighborhood poverty rates, and because it collected extensive data on teenagers, it provides us with the clearest answer so far to the threshold question of whether important neighborhood effects exist.

[38] Bloom, Riccio, and Verma (2005) found substantial positive earnings effects in a community-based intervention.
[39] Reviewing the early non-experimental literature, Mayer and Jencks (1989) conclude "the more we learn about a given outcome, the smaller the effects of mean SES usually look." More recently, Ellen and Turner (1997) report that "some recent studies that have done the most careful job of controlling for unobserved family characteristics . . . find no independent neighborhood effects, casting doubt on the robustness of results from other studies." Finally, recent quasi-experimental studies (Jacob, 2004; Oreopoulos, 2003) find little or no effect of living in high-poverty housing projects on child outcomes.

REFERENCES

Altonji, Joseph and Blank, Rebecca, 1999. "Race and Gender in the Labor Market." In *Handbook of Labor Economics, Vol. 3A*, edited by Orley Ashenfeleter and David Card. Amsterdam: North Holland.

Angrist, Joshua A.; Imbens, Guido W., and Rubin, Donald B., 1996. "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91, 444-472.

Arnott, Richard and John Rowse. "Peer Group Effects and Educational Attainment." *Journal of Public Economics* 32 (1987), 287-305.

Austen-Smith, David, and Fryer, Roland, 2005. "An Economic Analysis of Acting White." *Quarterly Journal of Economics* 120:2, 551-583.

De Bartolome, Charles. "Equilibrium and Inefficiency in a Community Model with Peer Effects." *Journal of Political Economy* 98 (February 1990), 100-113.

Becker, Gary and Kevin M. Murphy, 2000. *Social Economics*. Chicago: University of Chicago.

Benabou, Roland, 1993. "Workings of a City: Location, Education, and Production." *Quarterly Journal of Economics*, 108:3, 619-652.

Bloom, Howard S.,James A. Riccio, and Nandita Verma, 2005. *Promoting Work in Public Housing The Effectiveness of Jobs-Plus.* New York: MDRC.

Bos, Johannes; Brock, Thomas; Duncan, Greg; Granger, Robert; Huston, Aletha and Lloyd, Vonnie, 1999. *New Hope for People with Low Incomes: Two-Year Results of a Program to Reduce Poverty and Reform Welfare.* New York: MDRC.

Brock, William A. and Durlauf, Steven N., 2001. "Interactions-based Models." In James J. Heckman and Edward Leamer, eds., *Handbook of Econometrics, Volume 5*. Amsterdam: North-Holland, 3297-3380.

Brooks-Gunn, Jeanne, Greg J. Duncan, and Lawrence Aber, 1997. *Neighborhood Poverty: Context and Consequences for Children, Vol. I.* New York: Russell Sage Foundation.

Coleman, John and Leo B. Hendry, 1999. *The Nature of Adolescence.* London: Routledge.

Crane, Jonathan (1991). "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing." American Journal of Sociology 96 (5): 1226-1259.

Davidson, Russell, and MacKinnon, James G., 1993. *Estimation and Inference in Econometrics*, New York, Oxford University Press.

Ellen, Ingrid G. and Turner, Margery A., 1997. "Does Neighborhood Matter? Assessing the Recent Evidence." *Housing Policy Debate*, 8: 833-866.

Entwisle, Doris R.; Alexander, Karl L. and Olson, Linda S. (1994). "The Gender Gap in Math: Its Possible Origins in Neighborhood Effects." *Child Development* 67: 2400-2416.

Fernandez, Raquel and Richard Rogerson. "Income Distribution, Communities and the Quality of Public Education." *Quarterly Journal of Economics* 111, 135-164.

Fryer, Roland, and Torrelli, Paul, 2005. "An Empirical Analysis of Acting White." Unpublished manuscript, Harvard University.

Goering, John and Judith D. Feins, eds. (2003). *Choosing a Better Life: Evaluating the Moving to Opportunity Social Experiment*. Washington, DC: Urban Institute Press, 2003.

Granovetter, Mark. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83, 1420-1443.

Heckman, James J., 2001. "Accounting for Heterogeneity, Diversity and General Equilibrium in Evaluating Social Programs." *Economic Journal,* 111, 654-699.

Ihlanfeldt, Keith R. and Sjoquist, David J. 1998. "The Spatial Mismatch Hypothesis: A Review of Recent Studies and Their Implications for Welfare Reform." *Housing Policy Debate*, 9: 849-892.

Imbens, Guido W. and Angrist, Joshua D., 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica,* 62(2), 467-475.

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.

Jencks, Christopher, and Mayer, Susan E., 1990. "Residential Segregation and Black Job Opportunities." In *Inner City Poverty in the United States*, ed. by Larry Lynn Jr. and Michael McGeary. Washington, D.C.: National Academy Press, 187-222.

Jencks, Christopher, and Mayer, Susan E., 1990. "The Social Consequences of Growing Up in a Poor Neighborhood," In *Inner City Poverty in the United States*, ed. by Larry Lynn Jr. and Michael McGeary. Washington, D.C.: National Academy Press, 111-186.

Kain, John F. "Housing Segregation, Negro Employment and Metropolitan Decentralization." *Quarterly Journal of Economics*, 1968, 82, pp. 175-197.

Katz, Lawrence F.; Kling, Jeffrey R. and Liebman, Jeffrey B., 2001. "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment." *Quarterly Journal of Economics*, 116(2), 607-654.

Kawachi, Ichiro and Berkman, Lisa F., 2003. "Introduction," in Ichiro Kawachi and Lisa F. Berkman, eds., *Neighborhoods and Health.* Oxford: Oxford University Press, 1-19.

Kessler, Ronald C.; Barker, Peggy R.; Colpe, Lisa J.; Epstein, Joan F.; Gfroerer, Joseph C.; Hiripi, Eva; Howes, Mary J.; Normand, Sharon-Lise T.; Mandersheid, Ronald W.; Walters, Ellen E. and Zaslavsky, Alan M., 2003. "Screening for Serious Mental Illness in the General Population." *Archives of General Psychiatry*, 60, 184-189.

Kling, Jeffrey R., and Liebman, Jeffrey B., 2004. "Experimental Analysis of Neighborhood Effects on Youth." Princeton IRS Working Paper 483, March.

Kling, Jeffrey R.; Liebman, Jeffrey B. and Katz, Lawrence F., 2005. "Bullets Don't Got No Name: Consequences of Fear in the Ghetto," in Thomas S. Weisner, ed. *Discovering Successful Pathways in Children's Development: New Methods in the Study of Childhood and Family Life*. Chicago: University of Chicago Press, 243-281.

Kling, Jeffrey R.; Liebman, Jeffrey B.; Katz, Lawrence F., and Sanbonmatsu, Lisa, 2004. "Moving to Opportunity and Tranquility: Neighborhood Effects on Adult Economic Self-Sufficiency and Health from a Randomized Housing Voucher Experiment." Princeton IRS Working Paper 481, April.

Kling, Jeffrey R.; Ludwig, Jens and Katz, Lawrence F., 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.

Kraemer, Sebastian, 2000. "The Fragile Male," *British Medical Journal*, 321, 23-30.

Leventhal, Tama and Brooks-Gunn, Jeanne, 2003a. "Moving To Opportunity: an Experimental Study on Neighborhood Effects on Mental Health." *American Journal of Public Health*, 93(9), 1576-1582.

Leventhal, Tama and Brooks-Gunn, Jeanne, 2003b. "New York City Site Findings: The Early Impacts of Moving To Opportunity on Children and Youth," in John M. Goering and Judith D. Feins, eds., *Choosing a Better Life: Evaluating the Moving to Opportunity Social Experiment.* Washington, DC: The Urban Institute Press, 213-244.

Liebman, Jeffrey B.; Katz, Lawrence F., and Kling, Jeffrey R.., 2004. "Beyond Treatment Effects: Estimating the Relationship Between Neighborhood Poverty and Individual Outcomes in the MTO Experimental." Princeton IRS Working Paper 493, August.

Lucas, Jacqueline W.; Schiller, Jeannine S. and Benson, Veronica, 2004. *Vital and Health Statistics, Series 10, Number 218.* Washington, DC: National Center for Health Statistics.

Ludwig, Jens; Duncan Greg J. and Hirschfield, Paul, 2001. "Urban Poverty and Juvenile Crime: Evidence from a Randomized Housing-Mobility Experiment." *Quarterly Journal of Economics* 116 (2): 655-679.

Mallows, C. L., 1986. "Augmented Partial Residual Plots." *Technometrics* 28, 313-319.

Manski, Charles F., 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies,* 60, 531-542.

O'Brien, Peter C., 1984. "Procedures for Comparing Samples with Multiple Endpoints." *Biometrics*, 40, 1079-1087.

Olsen, Edgar O., 2000. "The Cost-Effectiveness of Alternative Methods of Delivering Housing Subsidies." Manuscript, University of Virginia.

Olsen, Edgar O., 2003. "Housing Programs for Low-Income Households," in Robert A. Moffitt, ed., *Means-Tested Transfer Programs in the Untied States*. Chicago: University of Chicago Press and NBER, 365-441.

Oreopoulos, Philip, 2003. "The Long-Run Consequences of Growing Up in a Poor Neighborhood," *Quarterly Journal of Economics*, 118(4), 1533-1575.

Orr, Larry; Feins, Judith D.; Jacob, Robin; Beecroft, Eric; Sanbonmatsu, Lisa; Katz, Lawrence F.; Liebman, Jeffrey B. and Kling, Jeffrey R., 2003. *Moving to Opportunity: Interim Impacts Evaluation.* Washington D.C.: U.S. Department of Housing and Urban Development.

Popkin, Susan J.; Harris, Laura E. and Cunningham, Mary K., 2002. *Families in Transition: A Qualitative Analysis of the MTO Experience*. Washington, DC: Urban Institute.

Ramirez_Valles, Jesus; Zimmerman, Marc A. and Juarez, Lucia (2002). "Gender Differences of Neighborhood and Social Control Process: A Study of the Timing of First Intercourse among Low-achieving, Urban, African American Youth." *Youth and Society* 33(3):418-442.

Sanbonmatsu, Lisa; Kling, Jeffrey R.; Duncan, Greg J.; Brooks-Gunn, Jeanne. 2004. "Neighborhoods and Academic Achievement: Results from the MTO Experiment." Princeton IRS Working Paper 492.

Walsh, B. et al. 2002. Placebo Response in Studies of Major Depression: Variable, Substantial, and Growing," *Journal of the American Medical Association*, 287: 1840-1847.

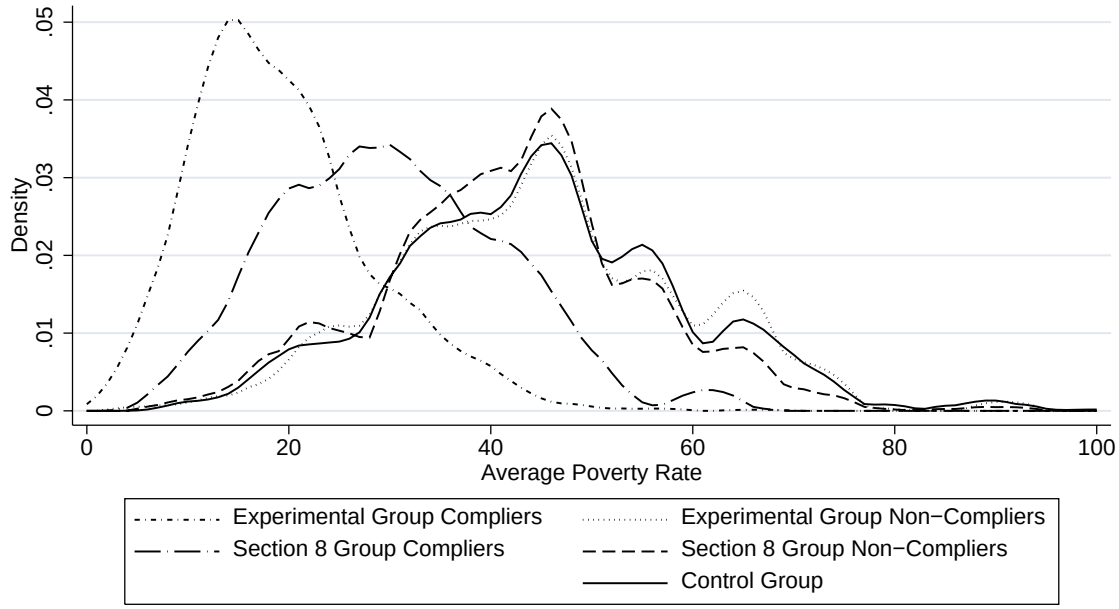
Wells, Kenneth B.; Sherbourne, Cathy; Schoenbaum, Michael; Duan, Naihua; Meredith, Lisa; Unützer, Jürgen; Miranda, Jeanne; Carney, Maureen F. and Rubenstein, Lisa V. 2000. "Impact of Disseminating Quality Improvement Programs for Depression in Managed Primary Care: A Randomized Controlled Trial," *Journal of the American Medical Association*, 283: 212-220.

Westfall, Peter H., and Young, S. Stanley, 1993. *Resampling-Based Multiple Testing: Examples and Methods for P-value Adjustment.* New York: Wiley.

Wilson, William J., 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy.* Chicago: University of Chicago Press.

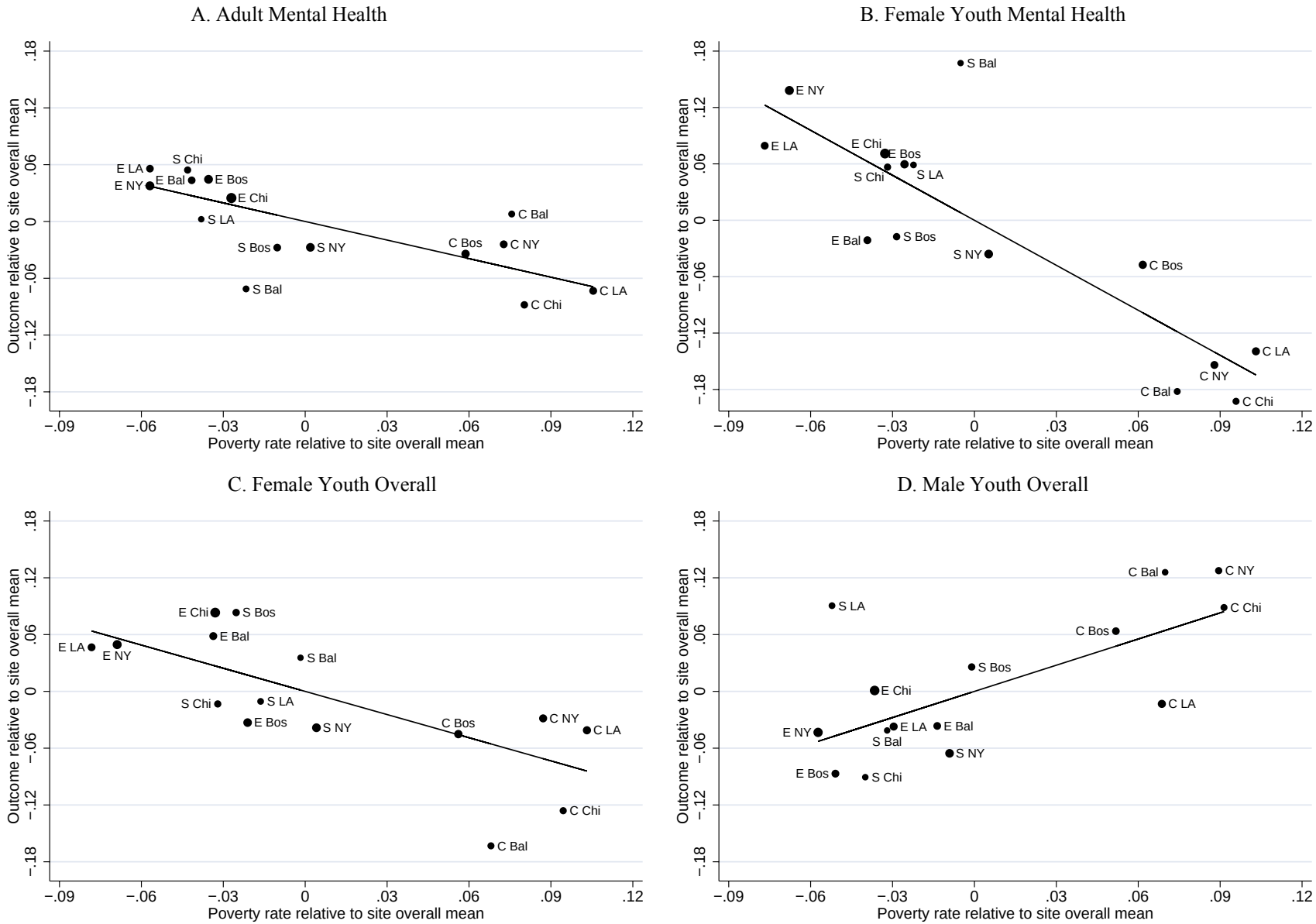
Zaslow, Martha J., and Hayes, Cheryl D., 1986. "Sex Differences in Child's Response to Psychosocial Stress: Toward a Cross-Context Analysis," in M. Lamb, A. Brown, and B. Rogoff, eds., *Advances in Developmental Psychology, Volume 4.* Hillsdale, NJ: Lawrence Earlbaum.

FIGURE 1. DENSITIES OF AVERAGE POVERTY RATE, BY GROUP



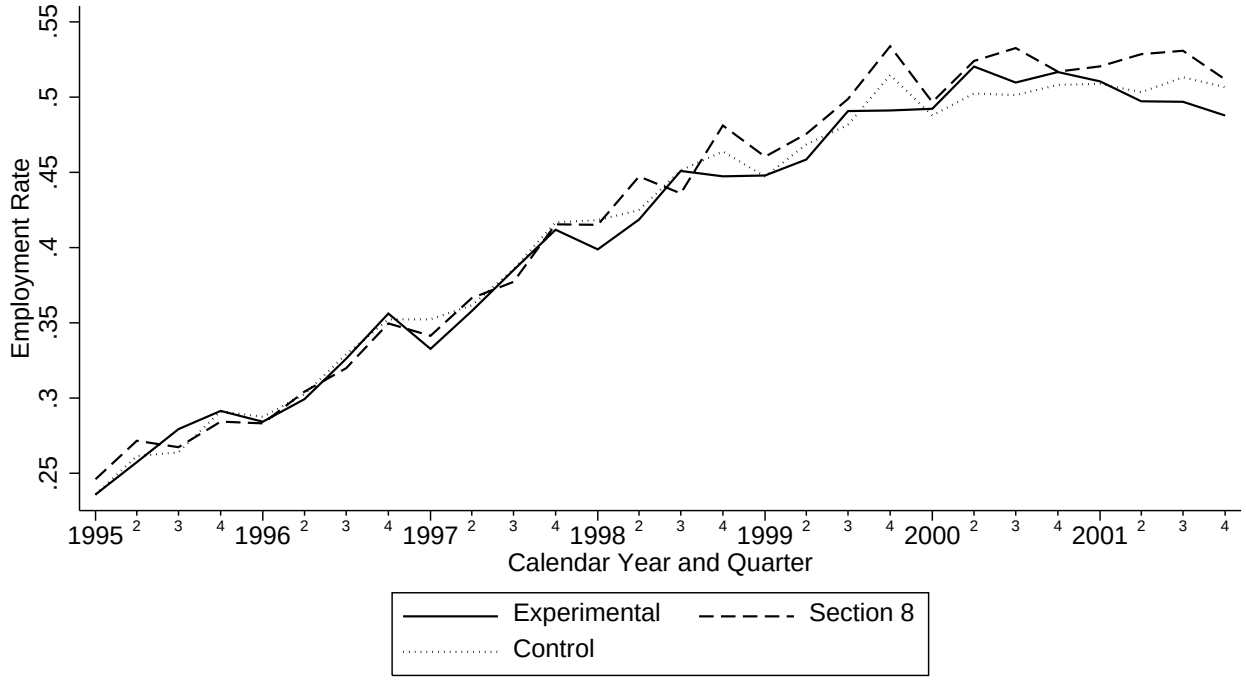
Notes. Average poverty rate is a duration-weighted average of tract locations from random assignment through 12/31/01. Poverty rate is based on linear interpolation of 1990 and 2000 Censuses. Density estimates using Epanechnikov kernel with halfwidth of 2.

FIGURE 2. PARTIAL REGRESSION LEVERAGE PLOTS



Notes. The index on the horizontal axis is expressed in standard deviation units relative to the control group overall standard deviation for each variable. The components of the overall and mental health indices are described in the notes to Table 2. The poverty rate is an average across tracts since random assignment, weighted by residential duration, using linear interpolation between the 1990 and 2000 Censuses. The line passes through the origin with the slope from 2SLS estimation of equation (3) of the outcome on poverty rate and site indicators, using group-by-site interactions as instrumental variables. The points are from a partial regression leverage plot of the group outcome means on the group poverty rate means, conditional on site main effects, as described in the text. The size of each point is proportional to the sample size of that group, and correspondingly to the weight each point receives in the 2SLS regression.

FIGURE 3. EMPLOYMENT RATES OVER TIME



Notes. Employment is fraction with positive earnings per quarter from Unemployment Insurance records in California, Illinois, Maryland, Massachusetts and New York.

TABLE 1. DESCRIPTIVE STATISTICS OF NEIGHBORHOOD CHARACTERISTICS

	Experimental	Section 8	Control
	(i)	(ii)	(iii)
Average census tract poverty rate	.33	.35	.45
Average census tract poverty rate above 30%	.52	.62	.87
Respondent saw illicit drugs being sold or used in neighborhood during past 30 days	.33	.34	.46
Streets are safe or very safe at night	.70	.65	.55
Member of household victimized by crime during past 6 months	.17	.16	.21
Average census tract share on public assistance	.16	.17	.23
Average census tract share of adults employed	.83	.83	.78
Average census tract share workers in professional and managerial occupations	.26	.23	.21
Average census tract share minority	.82	.87	.90

Notes. Census tract characteristics are the average for an individual's addresses from randomization through 2001 weighted by duration. Except for "managerial and professional occupations" (for which only 2000 Census data was used due to differences in the occupation classification used for the 1990 Census and 2000 Census), values for inter-census years are interpolated. "Saw illicit drugs," "streets are safe," and "victimized by crime" are based on adult report in 2002 survey. All Experimental - Control and Section 8 - Control differences have p-values < .05.

TABLE 2. MEAN EFFECT SIZES FOR SUMMARY MEASURES OF OUTCOMES

	All Adults		All Youth		Female Youth		Male Youth		M-F Youth	
	E-C	S-C	E-C	S-C	E-C	S-C	E-C	S-C	E-C	S-C
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Economic self-sufficiency	.017 (.031)	.037 (.033)								
Absence of physical health problems	.012 (.024)	.019 (.026)	-.038 (.038)	-.020 (.040)	.025 (.053)	.077 (.055)	-.112* (.053)	-.114* (.061)	-.138 (.076)	-.192* (.084)
Absence of mental health problems	.079* (.030)	.029 (.033)	.102 (.053)	.138* (.056)	.267* (.062)	.192* (.067)	-.052 (.080)	.054 (.092)	-.319* (.101)	-.138 (.113)
Absence of risky behavior			-.023 (.043)	-.039 (.050)	.142* (.053)	.129* (.059)	-.181* (.062)	-.208* (.071)	-.323* (.080)	-.337* (.092)
Education			.050 (.041)	.028 (.047)	.138* (.065)	.056 (.068)	-.053 (.047)	-.001 (.060)	-.191* (.080)	-.057 (.090)
Overall	.036 (.020)	.028 (.022)	.018 (.025)	.018 (.026)	.136* (.034)	.109* (.034)	-.099* (.031)	-.078* (.037)	-.235* (.047)	-.187* (.051)

Notes. E-C: Experimental - Control. S-C: Section 8 - Control. Estimates are the intent-to-treat mean effect sizes, from equation (1), fully interacted with gender in columns (v)-(x) as described in the text. The estimated equations all include site indicators and the baseline covariates listed in Appendix C. M-F Youth is male - female difference. Adult economic self-sufficiency: + adult not employed and not on TANF + employed + 2001 earnings - on TANF - 2001 government income. Adult mental health: - distress index - depression symptoms - worrying + calmness + sleep. Adult physical health: - self-reported health fair/poor - asthma attack past year - obesity - hypertension - trouble carrying/climbing. Adult overall includes 15 measures in self-sufficiency, physical health, and mental health. Youth physical health: - self-reported health fair/poor - asthma attack past year - obesity - non-sports injury past year. Youth mental health: - distress index - depression symptoms - anxiety symptoms. Youth risky behavior: - marijuana past 30 days - smoking past 30 days - alcohol past 30 days - ever pregnant or gotten someone pregnant. Youth education: + graduated high school or still in school + in school or working + WJ-R broad reading score + WJ-R broad math score. Youth overall includes 15 measures in physical health, mental health, risky behavior, and education. Sample sizes in the E, S, and C groups are 1453, 993, and 1080 for adults and 749, 510, and 548 for youth ages 15-20 on 12/31/01. Robust standard errors adjusted for household clustering are in parentheses.

TABLE 3. SPECIFIC OUTCOMES WITH EFFECTS SIGNIFICANT AT 5 PERCENT LEVEL

	E/S	CM	ITT	TOT	CCM
	(i)	(ii)	(iii)	(iv)	(v)
A. Adult outcomes					
Obese, BMI≥30	E-C	.468	-.051 (.022)	-.108 (.048)	.505
Calm and peaceful	E-C	.466	.061 (.022)	.131 (.047)	.443
Psychological distress, K6 z-score	E-C	.050	-.092 (.046)	-.196 (.099)	.150
B. Youth (female and male) outcomes					
Ever had generalized anxiety symptoms	E-C	.089	-.044 (.019)	-.099 (.042)	.164
Ever had generalized anxiety symptoms	S-C	.089	-.063 (.019)	-.138 (.055)	.147
Ever had depression symptoms	S-C	.121	-.039 (.019)	-.069 (.035)	.134
C. Female youth outcomes					
Psychological distress -- K6 scale z-score	E-C	.268	-.289 (.094)	-.586 (.197)	.634
Ever had generalized anxiety symptoms	E-C	.121	-.069 (.027)	-.138 (.055)	.207
Ever had generalized anxiety symptoms	S-C	.121	-.071 (.026)	-.125 (.046)	.168
Used marijuana in the past 30 days	E-C	.131	-.065 (.029)	-.130 (.059)	.202
Used marijuana in the past 30 days	S-C	.131	-.072 (.032)	-.124 (.056)	.209
Used alcohol in past 30 days	S-C	.206	-.091 (.038)	-.155 (.056)	.306
D. Male youth outcomes					
Serious non-sports accident or injury in past year	E-C	.062	.087 (.026)	.215 (.064)	0
Serious non-sports accident or injury in past year	S-C	.062	.080 (.028)	.157 (.058)	0
Ever had generalized anxiety symptoms	S-C	.055	-.049 (.024)	-.098 (.047)	.126
Smoked in past 30 days	E-C	.125	.103 (.032)	.257 (.084)	0
Smoked in past 30 days	S-C	.125	.151 (.037)	.293 (.073)	.014

Notes. E/S: indicates whether row is experimental - control (E-C) or Section 8 - control (S-C). CM: Control Mean. ITT: Intent-to-treat, from equation (1). TOT: Treatment-on-treated, from equation (2). CCM: Control complier mean. Robust standard errors adjusted for household clustering are in parentheses. The estimated equations all include site indicators and the baseline covariates listed in Appendix C. Rows shown in table to illustrate magnitudes were selected based on ITT p-values < .05, and are 17 of 60 from the set of specific outcomes (15 for adults and 15 for youth) and subgroups -- adults, youth (female and male), female youth, and male youth -- described in the notes to Table 2.

TABLE 4. EFFECTS OF NEIGHBORHOOD POVERTY RATES ON SELECTED OUTCOMES

Models	OLS	2SLS	2SLS	
RHS variables	Poverty (i)	Poverty (ii)	Poverty (iii)	Compliance (iv)
Adult Mental Health	.13 (.17)	-.62* (.24)	-1.35* (.60)	-.17 (.13)
Adult Overall	.16 (.12)	-.31* (.16)	-.53 (.39)	-.05 (.08)
Youth (female and male) Mental Health	.57 (.34)	-.97* (.41)	-.18 (.87)	-.20 (.21)
Female Youth Mental Health	.99 (.61)	-1.84* (.50)	-1.88 (1.09)	-.01 (.25)
Female Youth Risky Behavior	-.61 (.42)	-.94* (.39)	-1.03 (.85)	-.02 (.19)
Female Youth Overall	-.03 (.28)	-.90* (.26)	-1.03 (.56)	-.03 (.12)
Male Youth Physical Health	-.84* (.35)	1.07* (.49)	1.77 (1.09)	.18 (.26)
Male Youth Risky Behavior	-.06 (.42)	1.46* (.54)	.94 (1.29)	-.13 (.31)
Male Youth Overall	-.13 (.23)	.80* (.28)	1.47* (.68)	.17 (.16)

Notes. Models: OLS is from equation (3) with no excluded instruments, using the control group only; 2SLS is from equation (3) with 10 site-by-treatment interactions as excluded instruments, using the entire sample. Columns (i) and (ii) are each based on separate estimation of equation (3), with W including poverty rate. Each row in columns (iii) and (iv) contains coefficients from one estimate of equation (3) with W including poverty rate and an indicator for treatment compliance as endogenous variables. Units of summary indices are standard deviations of control group outcomes. The estimated equations all include site indicators and a full set of covariates combining baseline variables about adults and those about youth (for youth outcomes only): age, gender, race, marital status, employment, education, mobility history, attitudes about neighborhood, special classes for youth, behavioral or emotional problems of youth. Poverty rate averaged over tracts since random assignment, weighted by duration, using linear interpolation between 1990 and 2000 Censuses. Standard errors are in parentheses, adjusted for correlation between same-sex siblings. * = p-value <.05. Rows shown in table to illustrate magnitudes were selected based on 2SLS column (ii) p-value <.05, and are 9 of 19 from set of four adult, five youth (female and male), five female youth, and five male youth summary indices shown in Table 2.

Appendix

This appendix contains the following sections:

A. Summary indices and mean effect sizes

B. Calculation of adjusted p-values

C. Description of baseline covariates and outcomes

D. Comparison of outcomes to the National Longitudinal Survey of Youth

E. Additional discussion of internal validity

F. Additional results for adults

G. Additional results for youth

A. Summary indices and mean effect sizes

This paper reports results for outcomes that are summary indices, aggregating information across related outcomes. This aggregation improves statistical power to detect effects that are consistent across specific outcomes when these specific outcomes also have idiosyncratic variation. Focusing our interpretation on the indices helps us to form conclusions about the overall impact of the study and to reduce the number of statistical tests performed so as to reduce the chance of false positives. Specific outcomes are normalized to in standardized units to study mean effect sizes of the indices relative to the standard deviation of the control group.

To illustrate the creation of a summary index, the 15 specific outcomes for adults are shown in Table A1. Column 1 (labeled "raw") shows the mean of each outcome for the control group. In this paper, we focus on normalized transformations of each outcome (labeled "norm"), where we subtract the mean of the control group and divide by the standard deviation of the control group. Let Y_k be the *k*th of K outcomes, μ_k be the control group mean, and σ_k be the control group standard deviation. The normalized outcome is $Y_k^* = (Y_k - \mu_k)/\sigma_k$. The summary index is $Y^* = \Sigma_k Y_k^*/K$. We use the control group standard deviation to compare the treatment groups to their counterfactual, because this metric does not depend on which treatment (experimental or Section 8) is being analyzed.

In calculating the normed measure, we reverse the sign for adverse outcomes (welfare, government income, distress, depression, anxiety, poor general health, physical limitations, asthma, obesity, hypertension), so that a higher value of the normalized measure represents a more "beneficial" outcome. For earnings in 2001, the control group mean was 8829 and the experimental-control (E-C) difference was 246 -- a difference of .02 standard deviations, relative to the control group standard deviation. For asthma attack in the past year, the fraction having an attack was .21 in the control group, with an E-C difference of -.01. This is also a difference of .02 standard deviations, relative to the control group. This illustrates how we use this normalization in order to translate the magnitudes of different measures into standardized units.

The bottom row of Table A1 shows our summary index, which is the equally weighted average of the normalized transformations for each of the 15 outcomes. For twelve of the fifteen, the experimental group shows more beneficial outcomes than the control group, and the E-C difference for our summary index is .05 standard deviations. These results are based on unadjusted mean differences for simplicity of illustration, and slightly larger in magnitude (with

slightly smaller p-values) than our preferred regression-adjusted specification discussed in the text. For Table A1 and for the analyses in Table 2, weights are calculated based assuming that if an individuals was subsampled for any one outcome in the index then they were subsampled for all outcomes in the index.[40]

We interpret this summary index as aggregating information about related constructs, but do not intend to suggest that the measures within a domain are merely proxies for a single latent factor. For the fifteen key outcomes in our analysis of adults, there are three principal components with eigenvalues greater than one. Promax rotated factors do correspond to the a priori designation of the fifteen variables into the three domains pre-specified: economic self-sufficiency, physical health, and mental health. There is considerable variation that is not explained by the first principal component within each domain: 39% in economic self-sufficiency, 54% in mental health, and 64% in physical health. Relatedly, instead of equal weights of .2 on each variable, a principal components approach would have weights that ranged from .17-.24 in economic self-sufficiency, .12-.25 in mental health, and .13-.26 in physical health -- with lower weights on sleep, obesity, and hypertension. However, we do not believe that hypertension is less important than, say, asthma simply because it has lower correlation with self-reported overall health and with physical limitations (and consequently, with the first principal component of physical health); therefore, we do not adopt the principal component approach.

An alternative approach to estimating ITT effects on these summary constructs is to first estimate treatment effect for each outcome, standardized them, and then average them. This approach is very similar to that used for global significance testing in biostatistics (O'Brien 1984) and for effect sizes in educational meta-analysis (Hedges and Olkin 1985). Let σ_k^2 equal the variance of Y_k for the control group. Equation (4) defines the mean effect size, τ, for a set of K outcomes based on the treatment effect estimates and the control group standard deviations.

$$\textbf{(A1)} \quad \tau = \frac{1}{K}\sum_{k=1}^{K}\frac{\pi_{2k}}{\sigma_k}$$

[40] As discussed by Orr et al. (2003), subsampling was not conducted at the household level, but separately for youth surveys, testing, parental surveys, and blood pressure measurement depending upon what data had been collected at the time of subsampling. The assumption used in creating weights for indices is that an individual subsampled for any outcome (e.g., from the youth survey, testing, or parental report of high school completion) was subsampled for all outcomes in the index. This simplification drops data for a few individuals with partially complete information, but introduces no bias.

To calculate the sample variance of τ, we need to account for the covariance of the estimates π_{2k}. We obtain this covariance matrix using the seemingly unrelated regression system shown in equation (A2). Point estimates for each outcome are identical to those obtained using equation (1) for a specific outcome. Let I_K be a K by K identity matrix and let Z and X be defined as in (1).

$$\textbf{(A2)} \quad Y = \left(I_K \otimes \left(Z \quad X\right)\right)\theta + \upsilon \qquad\qquad Y = (Y_1^{'}, \ldots, Y_K^{'})'$$

We calculate a point estimate, standard error, and p-value for τ based on the parameters, π_{2k}, jointly estimated as elements of θ in (A2). These estimates treat σ_k as known. Kling and Liebman (2004) show that delta method and bootstrap approaches yield very similar inferences using these statistical methods in a study of MTO youth.

If there were no missing data on survey items and X contained only a constant, then the mean effect size in equation (A1) would be identical to estimation using the summary index in equation (1). Equation A2 is a more direct summary of the treatment effects on each specific outcome and it incorporates regression adjustment for each outcome. The summary index approach is simpler to compute and can be represented graphically which is why we use it in the paper.

A comparison of results from the two approaches is given in Table A2. As a practical matter, our results are not very sensitive to the specification for regression adjustment or to item nonresponse. Therefore, results from the two approaches are very similar.

Additional reference:

Hedges, Larry V. and Olkin, Ingram, 1985. *Statistical Methods for Meta-Analysis*. Orlando, FL: Academic Press.

Table A1. Components of Summary Indices for Adults

	CM		E-C		S-C	
	Raw 1	Norm 2	Raw 3	Norm 4	Raw 5	Norm 6
A. Economic Self-sufficiency						
Employed	.52	0	.02	.05	.02	.04
Earnings in 2001	8829	0	262	.02	-5	-.00
Employed & not on welfare	.45	0	.03	.06	.02	.04
Receiving welfare	.30	0	-.04*	.08*	-.05*	.11*
Government income in 2001	250	0	54	-.01	-158	.04
B. Physical health						
Overall health fair or poor	.33	0	.01	-.03	.02	-.03
Trouble carrying/climbing	.44	0	-.02	.04	-.02	.04
Asthma attack in past year	.21	0	-.01	.02	-.01	.02
Obese	.47	0	-.05*	.10*	-.05*	.09*
Hypertension	.30	0	.02	-.05	.03	-.06
C. Mental health						
Distress z-score	.05	0	-.09*	.09*	-.04	.04
Depression in past 12 months	.16	0	-.03*	.08*	-.02	.05
Worrying	.39	0	-.02	.05	-.01	.01
Calm and peaceful	.46	0	.07*	.13*	.02	.04
Sleep 7-8 hours nightly	.45	0	.04	.07	.02	.03
D. Adult overall index		0		.05*		.03

Notes. Raw = unadjusted value. Norm = (unadjusted value - control mean)/(control standard deviation); sign reversed for risky behavior, mental health, and physical health. CM = Control mean. E-C = Experimental - Control. S-C = Section 8 - Control. Differences based on unadjusted means, with no covariates. Summary index is the mean of normalized values of component items. Sample sizes are 859 and 3484 for male youth and all adults, respectively. * = p-value <.05.

Table A2. Summary Index and Mean Effect Size Results

	Summary Index		Mean Effect Size	
	E-C	S-C	E-C	S-C
	(i)	(ii)	(iii)	(iv)
Self-sufficiency	.017	.037	.016	.034
[5 measures]	(.031)	(.033)	(.031)	(.034)
Mental health	.079*	.029	.084*	.030
[5 measures]	(.030)	(.033)	(.030)	(.034)
Physical health	.012	.019	.016	.017
[5 measures]	(.024)	(.026)	(.024)	(.027)
Overall	.036	.028	.039	.027
[15 measures]	(.020)	(.022)	(.020)	(.022)

Notes. E-C: Experimental - Control. S-C: Section 8 - Control. Estimates are the mean of the standardized intent-to-treat effects, from equation (4). Standard errors are derived from equation (5), adjusted for correlation within individuals.

B. Calculation of adjusted p-values

This appendix describes our algorithm for calculating familywise adjusted p-values. It is based on the Westfall-Young (1993, algorithm 2.8) free step-down resampling method, modified to utilize per-comparison p-values based on bootstrap estimates instead of asymptotic approximations.

For each parameter of interest, τ_j, define $\hat{\tau}_j$ as the estimated value from the actual data and p^c_j as the asymptotic per-comparison p-value on the test of the null hypothesis that τ_j equals zero. Define N as the number of bootstrap replications. The per-comparison bootstrap p-value for τ_j is p^b_j, and the Westfall-Young familywise adjusted p-value for τ_j is p^a_j.

/* Calculate bootstrap p-values (p^b_j) */

For j = 1 to J {

    $p^a_j = p^b_j = 0$

}

For i = 1 to N {

    Draw a sample of households with replacement.

    For j = 1 to J {

        Calculate τ^*_{ij}, the estimated value of τ^*_j for this bootstrap replication.

        Calculate the p-value r_{ij} for the test that $\tau^*_{ij} = \hat{\tau}_j$.

        If $r_{ij} < p^c_j$, then $p^b_j = p^b_j + 1/N$

    }

}

/* Calculate p-values for each replication under null hypothesis (s_{ij}), ordering by r_{ij}
and imposing uniform p-value distribution across replications
for each of J parameters */

Define r_j as a vector of length N with elements r_{ij}

For j = 1 to J {

Sort elements of r_j so r_{kj} is smallest value of r_j when k is 1

For k = 1 to N {

$s_{kj} = (k-.5)/N$

}

}

/* Calculate adjusted p-value (p^a_j) */

For the J parameters in the family of tests, sort p^b_j such that j indexes family members in descending order of significance, so p^b_1 is the smallest bootstrap p-value.

For k = 1 to N {

$q_{J+1} = 1$

For j = J to 1 {

$q_j = \min(s_{kj} , q_{j+1})$

If $q_j < p^b_j$, then $p^a_j = p^a_j + 1/N$

}

}

/* Enforce monotonicity so that the order of outcomes according to bootstrap
per-comparison p-values is weakly preserved according to adjusted p-values */

$p^a_0 = 0$

For j = 1 to J {

$p^a_j = \max(p^a_{j-1} , p^a_j)$

}

C. Description of baseline covariates and outcomes

Covariates (*X)* used in equations (1) - (3) were drawn from data collected from a baseline survey conducted prior to random assignment. For analysis of adults, covariates included those in Table C1 and six Legendre polynomials for adult date of birth. For analysis of youth, all the covariates in Table C1 were used as well as those in Table C2, six Legendre polynomials for youth date of birth, and five indicators for missing data on: special class for gifted students or did advanced work; special school, class, or help for learning problem in past two years; special school, class, or help for behavioral or emotional problems in past two years; problems that made it difficult for him/her to get to school and/or to play active games or sports; suspended or expelled from school in past two years.

Descriptions of the 15 outcomes examined for adults and the 15 examined for youth are given below.

Adult economic self-sufficiency. Our measure of employment is an indicator for whether the adult had worked for pay during the week prior to survey.[41]

Welfare receipt is measured as being a beneficiary of Temporary Assistance for Needy Families (TANF) at the time of the survey.

Economic self-sufficiency is an indicator for working for pay during the previous week and not receiving TANF.

Earnings in 2001 is the amount self-reported to have been earned from all employers before taxes and deductions during 2001.

Government income is the amount received altogether in the form of TANF, Supplemental Security Income (SSI), unemployment benefits, Social Security, General Assistance and related programs in 2001.

Adult physical health. To assess overall health, respondents were asked "In general is your health excellent, very good, good, fair, or poor?" The analyses examined whether the respondent reported that he or she was in fair or poor health.

[41] At the time of our survey in 2002, MTO adults primarily worked in white-collar and service occupations, with hourly wages ranging from $7.44 at the 25th percentile to $12.47 at the 75th percentile. The data on occupation and hourly wage are taken from survey self-reports about the main job in the week of the interview. The majority of workers in white-collar occupations (46% of all employed) were cashiers, teacher's aides, secretaries, data-entry keyers, receptionists, typists, bookkeepers, adjusters, technicians, teachers, supervisors/proprietors/sales, or clerks. The majority of service workers (44% of all employed) were nursing aids, orderlies, janitors, maids, or guards. The three most common blue-collar occupations (10% of all employed) were taxicab driver/chauffeur, laborer, and packer.

Our measure of physical limitation was whether the respondent reported having at least a little trouble "lifting or carrying groceries" or "climbing several flights of stairs," two activities from traditional measures of activities of daily living likely to be relevant in a sample of mostly non-elderly adults (Ware et al., 1993).

Respondents were asked questions from the National Health Interview Survey sequence on asthma or wheezing attacks. As our dichotomous measure, we examined the fraction of respondents who had an attack during the past year.

Subjects self-reported their height and weight. We use the standard definition of obesity, BMI $\geq$ 30 kg/m^2 (National Institutes of Health 1998).

Our measure of hypertension is based on the JNC7 stage 1 systolic and diastolic blood pressure cut-points: systolic $\geq$ 140mm HG or diastolic $\geq$ 90mm HG (Chobanian et al., 2003). Systolic and diastolic blood pressures were measured from a single reading near the end of the survey from an Omron HEM-737. This device satisfied the American Association for the Advancement of Medical Instrumentation standards for accuracy (Anwar et al., 1998). Subjects were seated and had been at rest for at least 30 minutes.

Adult mental health. Distress during the past 30 days was assessed using the K6 scale, developed by Kessler et al. (2002). Additional psychometric analysis of the scale has been done by Furukawa et al. (2003). This scale score can range from 0 to 24, which we normalize to a z-score by subtracting the mean of 5.8 and dividing by the standard deviation of 5.4.

Depression was assessed using the Composite International Diagnostic Interview - Short Form (CIDI-SF; Kessler et al. 1998). If during a two-week period in the past year the respondent reported dysphoric mood (feeling "sad, blue or depressed") or anhedonia (having "lost interest in most things"), then he or she was assigned a probability of having had a major depressive episode (MDE) according to the number endorsed of seven possible symptoms corresponding to those used for Diagnostic and Statistical Manual of Mental Disorders - Fourth Edition (DSM-IV) psychiatric diagnosis. The probability is based on a mapping between the CIDI-SF screening questions and more detailed assessments in the National Comorbity Survey (Walters et al. 2003).[42]

[42] Version 1.0 of the World Health Organization's CIDI-SF contained a skip pattern error for people who reported dysphoric mood or anhedonia for "about half of the day" -- referred to here as "boundary cases. (CIDI-SF Memo: Edits, available at: http://www3.who.int/cidi/CIDI-SFeditsmemo.pdf. Accessed on July 3, 2003.) Due to the skip pattern error, boundary cases were not fully assessed to determine if the duration, intensity and frequency of their anhedonia met the minimum criteria for MDE. Assigning an MDE probability of zero to all boundary cases is one solution to this problem. We implement an alternative imputation that is also conservative but less extreme; we impute the probability for boundary cases of meeting the minimum criteria for MDE. For those with no indication

For worrying, respondents were asked the two initial screening items from the CIDI-SF sequence on generalized anxiety disorder, and we analyzed the fraction of the sample who answered "yes" to "felt worried, tense, or anxious" or "worried a lot more than most people would in your situation" (Kessler et al., 1998).

Respondents were asked if they felt "calm and peaceful" at least most of the time during the past month -- one of the items from the mental health inventory in the RAND Health Insurance Experiment and the Short Form-36 (SF-36; Ware et al., 1993).

Sleep was measured as the amount of time that participant usually spends sleeping each night, and we analyzed the fraction that usually sleep at least 7 and less than 9 hours per night. For discussion of the linkage between sleep and mental health, see Ford and Kamerow (1989).

Youth education. Our measures of whether a youth had graduated from high school or was still in school is based on a parental report.

To assess idleness, we asked youth whether they were in school, on summer vacation from school, working during the past week, or none of the above.

Reading achievement was assessed using the Woodcock-Johnson Broad Reading assessment, which includes letter-word identification and passage comprehension subtests.

Math achievement was assessed using the Woodcock Johnson Revised Broad Math assessment, which includes applied problems and calculation subtests. As discussed in Sanbonmatsu et al. (2004), some systematic differences in scores on tests administered by particular interviewers were detected for both reading and math tests, and the results presented in Table 3 are adjusted for potential interviewer effects after controlling for census tract fixed effects.

---

of anhedonia or anhedonia reported to be about half the day or less, we impute a probability of zero. For other cases with some indication of anhedonia that was not fully assessed, we reduce the probability of MDE caseness using a multiplicative factor of about one third to capture the probability that their anhedonia may not have been of sufficient duration or intensity. This imputation is conservative under the assumption that the probability of anhedonia with sufficient intensity and duration for MDE would be at least as high for cases with no dysphoric mood as for those with dysphoric mood for "about half of the day," because it is based on data from those reporting no dysphoric mood but who did report some anhedonia and who were fully assessed for its duration and intensity. Version 1.0 of the CIDI-SF also skipped symptom items if the respondent reported receiving medication for depression, with the intention of scoring these individuals as having a probability of one for MDE. With changes in prescription patterns since the CIDI-SF was developed a decade ago, the probability of MDE conditional on receiving medication is now less than one. We assume that those receiving medication had dysphoric mood or anhedonia and at least one other MDE symptom; since this sample has an overall probability of MDE of .81, we then impute the probability of MDE for those receiving medication to also be .81. Our analysis shows that the distributions of K6 distress scores were very similar for those receiving medication and for non-boundary cases with at least one symptom of MDE (whose symptoms were fully assessed) -- providing some support for the reasonableness of this procedure.

Youth physical health. Our analysis of youth health is based on self-reported information. We asked questions drawn mainly from the National Health Interview Survey about general health status, injuries, asthma attacks, height, and weight. Self-reported health is strongly related to life expectancy among adults (see Idler and Kasl 1995). While less is known about the predictive power of self-reported health in children, Case, Lubotsky, and Paxson (2002) and Currie and Stabile (2003) find that reported poor health correlates strongly with children's chronic conditions, bed days, and hospitalization episodes.

For asthma attacks, our measure follows the standard practice of combining attacks requiring medical attention with other episodes of wheezing or whistling in the chest (Pearce et al, 1998). The purpose of including the other episodes is to ameliorate potential confounding with health care access.

We asked for details of any injuries, accidents, or poisonings that required medical attention or were serious enough to limit activities during the previous twelve months, and we focused our analysis on non-sports injuries. We had hypothesized that non-sports injuries might decrease due to a reduction in dangerous external factors in treatment group neighborhoods, but that safety increases might be offset by greater sports participation and lead to no change or an increase in sports injuries.

To assess obesity, we collected self-reported height and weight and calculated the body mass index for each individual. Other studies that have collected self-reports and measurements indicate that older adolescents slightly over-report height and under-report weight and that the correlation between self-reported and direct measures is around .9 (Goodman et al. 2000, Brener et al. 2003). Our measure of obesity is body mass index greater than the 95^{th} percentile of the national norms for the youth's age and gender. In interpreting this measure, it is worth noting that national norms for height and weight are benchmarked to 1988-94 data and do not reflect the distribution in 2003, when the population appears to have been heavier.

Youth mental health. Our distress measure, developed by Kessler et al (2002) for the National Health Interview Survey, is commonly scored by summing the scale scores of the items, with the total ranging from 0-24; our results are reported as z-scores, scaling by the standard deviation. This measure is commonly known as the K6 and is based on a six-item Likert scale measuring how much of the time during the past 30 days the youth felt: "so depressed nothing could cheer you up," "nervous," "restless or fidgety," "hopeless," "everything was an effort," or "worthless." The sample mean is 5.0 and the standard deviation is 4.7.

Our measure of serious depression involved a series of screening questions about the duration and intensity of the feelings and the presence of related symptoms during the worst period in life. A youth is considered to have had a Major Depressive Episode during his or her lifetime if he or she met the following five conditions. A. The youth experienced a period in which for most of the day he or she felt one of the following: sad, empty or depressed, very discouraged or hopeless about how things were going in his or her life, or loss of interest and boredom with most things usually enjoyed like work, hobbies, and personal relationships. B. Either felt this way most of the day almost every day for a period of two weeks or longer, or for a period of three days or longer and had a year or more in life when felt this way just about every month for several days or longer. C. During times when mood was most severe and frequent, the feelings usually lasted not less than 3 hours a day. D. These feelings were either more than mild, sometimes felt so bad that nothing could cheer him or her up, or sometimes felt so bad that he or she could not carry out daily activities. E. These feelings were accompanied by changes in sleeping, eating, energy, his or her ability to keep mind on things, feeling badly about his or herself, or other problems.

Our measures of anxiety also involve a series of screening questions about the duration and intensity of the feelings and the presence of related symptoms during the worst period in life. A youth was considered to have had Generalized Anxiety Disorder during his or her lifetime if the following four criteria were met: A. The youth reported there was a period when he or she was either worrying a lot more about things than other people with the same problems, much more nervous or anxious than most people with the same problems, or anxious or worried most days. B. The youth reported being worried about nothing in particular, everything, or more than one specific thing. C. The youth sometimes or often either found it hard to stop the worries or anxiety or could not think about anything else no matter how hard he or she tried. D. The period of being anxious, nervous, or worried lasted at least one month. Depression and anxiety were chosen because they are sufficiently common in the population to ensure that their minimum detectable effects were reasonable with our sample sizes. Since these serious events are still relatively rare, we examined the prevalence of any event, rather than focusing on a limited time such as the past year. Some of the reported events occurred prior to random assignment; we did not attempt to date them precisely, but instead rely on the assumption that these were few in number (since prevalence of reported events is extremely low at young ages) and that the prevalence of these early events was similar on average between the randomly assigned groups.

Youth risky behavior. Our measures of risky behavior drew upon survey self-reports using items from the National Longitudinal Survey of Youth 1997.

Alcohol use was measured as any use within the past 30 days.

Smoking was measured as any cigarette smoking within the past 30 days.

Our measures of whether a female youth had borne a child or a male youth had fathered a child were also based on self-reported responses to survey questions.

Additional references:

Anwar, Yusra Anis; Giacco, Sharon; McCabe, Ellen J.; Tendler, Beatriz E. and White, William B., 1998. "Evaluation of the Efficacy of the Omron HEM-737 Intellisense Device For Use On Adults According to the Recommendations of the Association for Advancement of Medical Instrumentation." *Blood Pressure Monitoring*, 3, 261-265.

Brener, Nancy D.; McManus, Tim; Galuska, Deborah A.; Lowry, Richard and Wechsler, Howell, 2003. "Reliability and Validity of Self-reported Height and Weight Among High School Students." *Journal of Adolescent Health* 32: 281-287.

Case, Anne C.; Lubotsky, Darren and Paxson, Christine H., 2002. "Economic Status and Health in Childhood: The Origins of the Gradient." *American Economic Review* 92 (5): 1308-1334.

Chobanian, Aram V.; Bakris, George L.; Black, Henry R.; Cushman, William C.; Green, Lee A.; Izzo, Joseph L. Jr.; Jones, Daniel W.; Materson, Barry J.; Oparil, Suzanne; Wright, Jackson T. Jr. and Roccella, Edward J., 2003. "The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report." *Journal of the American Medical Association,* 289, 2560-2572.

Currie, Janet and Stabile, Mark, 2003. "Socioeconomic Status and Health: Why is the Relationship Stronger for Older Children?" *American Economic Review* 93(5), 1813-1823.

Furukawa, Toshiaki A.; Kessler, Ronald C.; Slade, T. and Andrews, G., 2003. "The Performance of the K6 and K10 Screening Scales for Psychological Distress in the Australian National Survey of Mental Health and Well-Being." *Psychological Medicine*, 33, 357-362.

Goodman, Elizabeth and Whitaker, Robert C., 2002. "A Prospective Study of the Role of Depression in the Development and Persistence of Adolescent Obesity." *Pediatrics* 109: 497-504.

Idler, Ellen L. and Kasl, Stanislav V. (1995). "Self-Ratings of Health: Do They Also Predict Changes in Functionality." *Journal of Gerontology: Social Sciences* 50 (6): S344-353.

Kessler, Ronald C.; Andrews G.; Mroczek, D.K.; Ustun. T.B. and Wittchen, H.U., 1998. "The World Health Organization Composite International Diagnostic Interview Short-Form (CIDI-SF)." *International Journal of Methods in Psychiatric Research,* 7, 171-185.

Kessler, Ronald C.; Andrews, G.; Colpe, Lisa J.; Hiripi, Eva; Mroczek, Daniel K.; Normand, Sharon-Lise T.; Walters, Ellen E. and Zaslavsky, Alan M., 2002. "Short Screening Scales to Monitor Population Prevalences and Trends in Non-specific Psychological Distress." *Psychological Medicine*, 32, 959-976.

National Institutes of Health, 1998. "Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults--The Evidence Report." *Obesity Research,* 6(Supp 2), 51S-209S.

Pearce, Neal; Beasley, Richard; Burgess, Carl and Crane, Julian, 1998. *Asthma Epidemiology: Principles and Methods*. Oxford: Oxford University Press.

Walters Ellen E.; Kessler, Ronald C.; Nelson Christopher B. and Mroczek, Daniel M., 2002. "Scoring the World Health Organization's Composite International Diagnostic Interview Short Form (CIDI-SF)." (Accessed 10/26/04; http://www3.who.int/cidi/CIDISFScoringMemo12-03-02.pdf)

Ware, John E.; Snow, K.K.; Kosinski, M. and Gandek, B., 1993 *SF-36 Health Survey: Manual and Interpretation Guide.* Boston: The Health Institute, New England Medical Center.

Table C1. Adult Baseline Characteristics

Variable	Control	Experimental				Section 8			
	Mean	Mean	CP Mean	NCP Mean	CP-NCP	Mean	CP Mean	NCP Mean	CP-NCP
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
Demographics									
Age in years (as of December 2001)	39.6	39.7	38.1	41.2	-3.1*	40.1	38.3	42.6	-4.2*
Male	.02	.01	.00	.02	-.01*	.02	.02	.02	.00
Baltimore site	.15	.15	.17	.13	.04	.15	.18	.10	.09*
Boston site	.21	.22	.21	.22	-.02	.22	.18	.28	-.10*
Chicago site	.22	.23	.16	.30	-.14*	.23	.25	.19	.06
Los Angeles site	.16	.16	.21	.10	.11*	.15	.19	.09	.11*
New York site	.25	.25	.25	.24	.01	.25	.19	.34	-.15*
African-American	.66	.67	.67	.66	.01	.66	.70	.60	.10*
Other race	.27	.26	.23	.29	-.05*	.26	.22	.31	-.09*
Hispanic ethnicity, any race	.29	.29	.28	.29	-.02	.30	.27	.35	-.08*
Never married	.62	.62	.66	.58	.08*	.62	.65	.58	.07
Teen parent	.24	.25	.26	.24	.02	.26	.30	.21	.08*
Economic and education									
Working	.25	.29	.29	.28	.02	.25	.26	.24	.03
On AFDC	.75	.74	.76	.72	.04	.75	.78	.70	.08*
In school	.16	.16	.20	.12	.07*	.16	.18	.12	.05*
High school diploma	.38	.41	.41	.42	-.01	.41	.41	.40	.01
General equivalency diploma	.21	.18	.21	.15	.06*	.19	.20	.18	.01
Household									
Had car	.15	.17	.19	.15	.04	.16	.18	.14	.05
Household member with a disability	.16	.16	.15	.17	-.02	.17	.14	.20	-.06*
Household member victimized by crime during past 6 months	.41	.42	.46	.39	.07*	.43	.45	.39	.05
No teen children	.62	.59	.66	.53	.13*	.61	.67	.52	.15*
Household of size 2	.20	.23	.27	.19	.09*	.21	.23	.18	.05
Household of size 3	.32	.30	.31	.30	.01	.31	.30	.31	-.01
Household of size 4	.22	.23	.23	.24	-.01	.23	.23	.22	.00
Neighborhood and housing									
Lived in neighborhood 5 or more years	.62	.61	.60	.62	-.02	.63	.57	.72	-.15*
Moved more than 3 times in past 5 years	.11	.08+	.09	.07	.02	.09	.11	.06	.05*
Very dissatisfied with neighborhood	.46	.46	.52	.41	.11*	.47	.52	.39	.13*
Streets very unsafe at night	.49	.48	.52	.45	.07*	.49	.53	.43	.10*

Table C1. Adult Baseline Characteristics, continued

Variable	Control	Experimental				Section 8			
	Mean	Mean	CP Mean	NCP Mean	CP-NCP	Mean	CP Mean	NCP Mean	CP-NCP
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
Neighborhood and housing									
Lived in neighborhood 5 or more years	.62	.61	.60	.62	-.02	.63	.57	.72	-.15*
Moved more than 3 times in past 5 years	.11	.08+	.09	.07	.02	.09	.11	.06	.05*
Very dissatisfied with neighborhood	.46	.46	.52	.41	.11*	.47	.52	.39	.13*
Streets very unsafe at night	.49	.48	.52	.45	.07*	.49	.53	.43	.10*
Chats with neighbors at least once a week	.55	.52	.49	.55	-.06*	.50	.50	.50	.00
Respondent very likely to tell neighbor if saw neighbor's child getting into trouble	.56	.53	.50	.57	-.07*	.55	.56	.53	.03
No family living in neighborhood	.65	.65	.66	.64	.02	.62	.63	.60	.03
No friends living in neighborhood	.41	.40	.43	.38	.05	.38	.40	.34	.06
Very sure would find an apartment in another part of city	.45	.45	.51	.40	.11*	.48	.54	.40	.14*
To get away from gangs or drugs was primary or secondary reason for moving	.78	.77	.79	.75	.04	.75	.77	.73	.05
Better schools was primary or secondary reason for moving	.48	.47	.50	.46	.04	.52	.53	.49	.05
Had applied for Section 8 voucher before	.45	.41	.44	.39	.05	.39+	.38	.40	-.03
N	1080	1453	694	759		993	585	408	

Notes: AFDC = Aid to Families with Dependent Children. CP = complier; NCP = non-complier. * = difference between treatment compliers and non-compliers is statistically significant at the 5 percent level. + = difference between treatment and control mean is statistically significant at 5 percent level.

Table C2. Youth Baseline Characteristics

	Female			Male		
	Exp (1)	Sec8 (2)	Con (3)	Exp (4)	Sec8 (5)	Con (6)
African-American	.68	.64	.67	.64	.65	.59
Special class for gifted students or did advanced work	.15	.17	.17	.17*	.15*	.27
Special school, class, or help for learning problem in past two years	.13	.13	.12	.29	.25	.30
Special school, class, or help for behavioral or emotional problems in past two years	.07	.08	.05	.18	.17	.11
Problems that made it difficult to get to school and/or to play active games	.03	.06	.06	.11*	.08	.05
Problems that required special medicine and/or equipment	.05	.07	.05	.13	.14	.09
School asked to talk about problems child having with schoolwork or behavior in past two years	.19	.23	.19	.41	.37	.33
Suspended or expelled from school in past two years	.09	.10	.07	.23	.20	.15

Notes. Exp: Experimental. Sec8: Section 8. Con: Control. * indicates p-value <.05 on difference between experimental or Section 8 and control group. Baseline data was collected at random assignment, during 1994-1997. Surveys were completed in experimental, Section 8 and control groups with 749, 510, and 548 respondents respectively ages 15-20 on 12/31/2001 for a total sample size of 1807.

D. Comparison of outcomes to the National Longitudinal Survey of Youth

Table D1. Risky behavior outcome means

	MTO			NLSY97	
	Exp	Sec8	Control	adjusted	unadjusted
	(1)	(2)	(3)	(4)	(5)
A. Females					
Used marijuana in past 30 days	.07	.08	.13	.13	.16
Smoked in past 30 days	.14	.14	.19	.25	.33
Had alcohol in past 30 days	.14	.14	.21	.28	.44
Been or gotten someone pregnant	.25	.33	.27	.21	.14
B. Females -- gifted dropped					
Used marijuana in past 30 days	.07	.08	.14	.12	.15
Smoked in past 30 days	.14	.11	.20	.26	.33
Had alcohol in past 30 days	.13	.12	.23	.28	.43
Been or gotten someone pregnant	.26	.32	.27	.22	.14
C. Males					
Used marijuana in past 30 days	.19	.21	.12	.21	.18
Smoked in past 30 days	.24	.29	.13	.29	.33
Had alcohol in past 30 days	.21	.24	.14	.30	.46
Been or gotten someone pregnant	.16	.19	.12	.16	.07
D. Males -- gifted dropped					
Used marijuana in past 30 days	.18	.23	.13	.21	.18
Smoked in past 30 days	.22	.30	.12	.29	.33
Had alcohol in past 30 days	.19	.22	.13	.30	.45
Been or gotten someone pregnant	.17	.17	.12	.16	.07

Notes. Exp: Experimental. Sec8: Section 8. Con: Control. Columns 1-3 are unadjusted means using MTO survey weights. Column 5 is the unadjusted sample mean of NLSY97 Round 3 outcomes for ages 15-20 using NLSY97 survey weights. Using the same NLSY97 data, column 4 contains the predicted values from regressions of outcomes on covariates, based on MTO covariate means. Covariates were sixth order polynomial in age, race white, race other non-black, adult head age 19-29, adult head age 30-39, adult head age 40-49, household size 2, household size 3, household size 4, adult head has car, adult head employed, adult head GED or high school graduate, adult head receiving welfare, missing parental interview, youth gifted classes, youth remedial classes, youth disabled, youth special medical needs. MTO covariates are from the MTO baseline survey. NLSY97 age is as of Round 3 interview; other NLSY covariates are from Round 1, recoded to match MTO baseline covariates. Regressions were estimated separately for females and males and evaluated at the gender-specific means of the MTO baseline covariates (except missing parental interview indicator evaluated at NLSY97 mean). Panels B and D drop observations where youth had earlier been in gifted classes to illustrate the lack of sensitivity to the covariate Imbalance shown in Table C2.

E. Additional discussion of internal validity

Regarding the internal validity of these results, two key concerns include the use of self-reports and the possibility of attrition bias. Most of the outcomes used in this paper were self-reported, and neither the participants nor the interviewers were blinded to the intervention. Thus, it is possible that the estimated impacts are due to some sort of reporting bias. However, the consistency between survey and administrative self-sufficiency estimates discussed in the main text and the negligible estimates of treatment effects for many outcomes help rule out the most obvious types of reporting bias. Given that the name of the demonstration is "Moving to Opportunity" and that it was promoted by HUD as a pathway to better jobs, one might expect employment and earnings to be the most likely outcomes to be exaggerated by the treatment groups, but this did not turn out to be the case. Also, in related MTO research studying youth arrests (Kling, Ludwig, and Katz 2005), self-reported and administrative data have generated similar results.

Additional supporting evidence finds strong beneficial effects on the mental health of female youth (Kling and Liebman 2004). Female adults and youth may have experienced similar outcomes from living in the same neighborhood, and the youth tend to have less awareness that their household had been randomly assigned to a group in the MTO demonstration five years ago and seem even less likely to provide biased reports.

Because some participants and interviewers were aware of treatment status, it is possible that some survey responses reflected what the participants thought the investigators wanted to hear rather than the truth or that interviewers themselves (though not told whether a respondent was a member of the intervention group) might surmise which group the respondent was in from where the person lives and somehow administer the questions or record the answers differently. If respondents were giving positive responses because they "won the lottery," then we would have expected the Section 8 group (which received the most desirable lottery outcome, an unrestricted voucher) to report more positive responses than the Experimental group (which received a geographically restricted voucher) – but this did not occur for any outcome. For social desirability bias to be consistent with the results for youth, it would have to be very complex – positive bias for female substance use and mental health, negligible for female physical health, and negative for males – and the available evidence is not consistent with a broad, systematic effect of this sort. On measures where one might expect a strong social

desirability bias, such as obesity, poor health, dropping out of high school, or being idle (not working or in school), there are not significant treatment effects for youth. Moreover, using the same type of demographic adjustments as in Table C1, we find that the MTO treatment groups are within a couple of percentage points of similar youth in the NLSY97 on these measures (see Table D1), whereas social desirability bias might predict that they would report significantly more desirable behavior. A lack of systematic social desirability bias between the treatment and control groups is consistent with a low level of awareness among youth about treatment status from a housing voucher lottery that their parents participated in when they were ages 8-16 and how it affected their residential location 4-7 years later when they were 15-20 years old. To the extent that outcomes like risky behavior are under-reported by a constant factor (say, two-thirds of the time) in all groups, the lower prevalence in self-reported data does reduce the statistical power to detect treatment effects, but does not bias their direction or result in the appearance of treatment effects when the true effects are zero.

In terms of potential attrition bias, our effective survey response rate was 90 percent and it is possible that the characteristics of those who were not interviewed differed systematically across the three groups. However, response rates were similar across the randomly-assigned groups, and our estimation models control for baseline characteristics in order to reduce the sensitivity of our results to differential attrition. Of course, it is also possible that the individuals who were not interviewed in the three groups differed in their unobservable characteristics. Kling and Liebman (2004) conduct extensive bounds calculations for youth outcomes from the MTO interim evaluation. They show that worse case assumptions about missing data can change the results a great deal, but that the sign of summary measure estimates do not change under less extreme assumptions about missing data.

We have used the administrative data on employment, earnings, and welfare to compare estimates for full sample and for the sample with which we completed surveys and did not find significant differences. This analysis was based on the four states with individual-level UI data, and the five states with individual-level welfare data. For example, the experimental group ITT estimate of the five years after RA was .024 for employment and -.017 for welfare in the full sample, and .038 for employment and -.022 for welfare in the sample with completed surveys (using survey weights), with p-values on the differences of .25 for employment and .62 for welfare. The point estimates of the employment rates for the survey sample were consistently

higher than for the full administrative sample, and the p-values on this contrast for the six employment and earnings measures in Table 4 ranged from .40 to .12. Further comparisons of the full sample to everyone we attempted to interview regardless of completion status (and therefore involving no attrition) found differences just as large or larger. Thus even these modest and statistically insignificant differences seem more likely to be the result of sampling variation from our subsampling of nonrespondents than of attrition bias.

F. Additional results for adults

Table F1. List of Tables for Adult Results

Table	Title
F2	Effects on Selected Mediating Factors
F3	Effects on Economic Self-Sufficiency
F4	Effects on Earnings and Welfare Receipt ─ Administrative Data
F5	Effects on Mental Health and Physical Health
F6	Effects on Economic Self-Sufficiency and Health by Age at Randomization
F7	Effects on Employment by Age at Randomization, Administrative Data
F8	Effects on Earnings by Age at Randomization, Administrative Data
F9	Effects on Voucher Use, Housing and Neighborhood Quality, and Safety
F10	Effects on Social Networks
F11	Effects on Education, Training, Health Behaviors and Health Care Access
F12	Effects on Mobility and Housing Assistance, Access to Transportation, and Relative Income
F13	Baseline Characteristics of Adult Survey Respondents and the Full Adult Sample
F14	Effects on Change in Employment within Zip Code Between 1994 and 2001

Table F2. Effects on Selected Mediating Factors

		Experimental versus Control				Section 8 versus Control			
	CM	ITT	TOT	CCM	N	ITT	TOT	CCM	N
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
Average census tract poverty rate [ADDR]	.448	-.119* (.007)	-.256* (.012)	.449	2533	-.097* (.006)	-.160* (.010)	.463	2073
Average census tract poverty rate below 30% [ADDR]	.132	.345* (.018)	.739* (.031)	.131	2533	.242* (.020)	.401* (.031)	.130	2073
Average census tract share on public assistance [ADDR]	.228	-.063* (.004)	-.136* (.008)	.227	2533	-.055* (.004)	-.091* (.006)	.239	2073
Average census tract share of adults employed [ADDR]	.384	.074* (.004)	.159* (.008)	.386	2532	.056* (.004)	.093* (.006)	.379	2072
Average census tract share workers in professional and managerial occupations [ADDR]	.215	.041* (.004)	.087* (.008)	.207	2530	.016* (.004)	.027* (.007)	.210	2071
Respondent saw illicit drugs being sold or used in neighborhood during past 30 days [SR]	.457	-.118* (.022)	-.253* (.046)	.432	2481	-.104* (.024)	-.171* (.039)	.451	2023
Average census tract share minority [ADDR]	.898	-.074* (.007)	-.159* (.014)	.886	2533	-.025* (.007)	-.042* (.012)	.896	2073
Average census tract share minority below 50% [ADDR]	.058	.065* (.011)	.140* (.024)	.064	2533	.006 (.010)	.010 (.017)	.062	2073
Moved at least 10 miles from baseline address [ADDR]	.106	.054* (.016)	.116* (.034)	.154	2424	.028 (.018)	.046 (.030)	.111	2005
Housing has problem with mice, rats or cockroaches [SR]	.541	-.049* (.022)	-.104* (.046)	.479	2511	-.014 (.023)	-.024 (.039)	.500	2058
Has a friend who graduated college or who earns more than $30,000 a year [SR]	.518	.053* (.022)	.113* (.047)	.513	2334	.032 (.025)	.054 (.042)	.511	1917
Attends church or religious service at least once a month [SR]	.426	-.031 (.021)	-.066 (.046)	.464	2521	.008 (.024)	.014 (.039)	.438	2064

Notes. ADDR = address history from tracking file linked to Census data. Census tract characteristics are the average for an individual's addresses from randomization through 2001 weighted by duration. Except for "managerial and professional occupations" (for which only 2000 Census data was used due to differences in the occupation classification used for the 1990 Census and 2000 Census), values for inter-census years are interpolated. SR = self-report. CM = control mean. Intent-to-treat (ITT) from equation (1), using covariates in Table C1 and weights described in the text. TOT = Treatment-on-treated from equation (2) estimated by two stage least squares with treatment group assignment indicator variables as the instruments for the treatment take-up indicator variables. CCM = control complier mean, as defined in the text. * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses.

Table F3. Effects on Economic Self-Sufficiency

		Experimental versus Control				Section 8 versus Control			
	CM	ITT	TOT	CCM	N	ITT	TOT	CCM	N
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
Adult employed and not on TANF [SR]	.453	.019 (.020)	.040 (.044)	.453	2521	.015 (.023)	.025 (.038)	.449	2066
Employed [SR]	.520	.015 (.021)	.033 (.044)	.533	2525	.024 (.023)	.040 (.038)	.522	2068
Earnings in 2001 [SR]	8839	125 (449)	268 (960)	9108	2386	-5 (486)	-9 (811)	9305	1950
Receiving TANF [SR]	.295	-.021 (.019)	-.046 (.040)	.325	2519	-.031 (.021)	-.051 (.034)	.320	2063
Income received from government sources during 2001 [SR]	2484	194 (184)	419 (398)	2248	2381	-110 (205)	-181 (336)	2297	1946

Notes. TANF = Temporary Assistance for Needy Families. SR = self-report. CM = control mean. Intent-to-treat (ITT) from equation (1), using covariates in Table C1 and weights described in the text. TOT = Treatment-on-treated from equation (2) estimated by two stage least squares with treatment group assignment indicator variables as the instruments for the treatment take-up indicator variables. CCM = control complier mean, as defined in the text. * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses.

Table F4. Effects on Earnings and Welfare Receipt ─ Administrative Data

		Experimental versus Control				Section 8 versus Control			
	CM	ITT	TOT	CCM	N	ITT	TOT	CCM	N
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
A. Employment									
Fraction of quarters employed in 2001 [ADMIN]	.508	-.017 (.017)	-.036 (.035)	.550	2910	.014 (.017)	.022 (.028)	.546	2411
Fraction of quarters employed in years 1 through 5 after RA [ADMIN]	.422	-.006 (.013)	-.012 (.028)	.468	2455	.001 (.014)	.001 (.023)	.447	2039
Fraction of quarters employed in year 5 after RA [ADMIN]	.499	.002 (.018)	.005 (.039)	.532	2455	.008 (.020)	.013 (.032)	.531	2039
B. Earnings									
Earnings in 2001 [ADMIN]	8490	-287 (400)	-612 (853)	9062	2910	41 (441)	67 (714)	8899	2411
Annualized earnings in years 1 through 5 after RA [ADMIN]	5948	-6 (295)	-13 (630)	5622	2455	90 (345)	143 (546)	5481	2039
Earnings in year 5 after RA [ADMIN]	7924	128 (417)	273 (890)	7475	2455	370 (471)	587 (744)	7313	2039
C. TANF receipt									
Fraction of quarters received TANF in 2001 [ADMIN]	.263	-.001 (.015)	-.001 (.031)	.281	2912	.005 (.016)	.008 (.026)	.265	2407
Fraction of quarters received TANF in year 5 after RA [ADMIN]	.276	-.011 (.018)	-.024 (.040)	.293	2041	.018 (.021)	.029 (.033)	.264	1569
D. TANF amount									
Amount of TANF received in 2001 [ADMIN]	1406	-44 (88)	-92 (187)	1653	2912	-92 (94)	-150 (153)	1493	2407
Amount of TANF payments received in year 5 after RA [ADMIN]	1316	-116 (96)	-263 (219)	1500	2041	7 (110)	11 (176)	1242	1569

Notes. ADMIN = administrative data. RA = random assignment. TANF = Temporary Assistance for Needy Families. CM = control mean. Intent-to-treat (ITT) from equation (1), using covariates in Table C1 and weights described in the text. TOT = Treatment-on-treated from equation (2) estimated by two stage least squares with treatment group assignment indicator variables as the instruments for the treatment take-up indicator variables. CCM = control complier mean, as defined in the text. Administrative data on employment and earnings are from state unemployment insurance (UI) records and data on TANF receipt are from state and county welfare agencies. Data were obtained for California (LA county only for TANF), Illinois, Maryland, Massachusetts and New York. TANF data were analyzed at the individual level. UI estimates are based on cell data as described in the text, controlling for site and mean randomization quarter, baseline education, and baseline work status. * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses.

Table F5. Effects on Mental Health and Physical Health

		Experimental versus Control				Section 8 versus Control			
	CM	ITT	TOT	CCM	N	ITT	TOT	CCM	N
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
A. Mental health									
Psychological distress, K6 z-score [SR]	.050	-.092* (.046)	-.196* (.099)	.150	2531	-.033 (.051)	-.054 (.085)	.028	2069
Probability of major depressive episode [SR]	.164	-.027 (.014)	-.059 (.031)	.196	2529	-.013 (.016)	-.022 (.027)	.165	2070
Worried, tense, or anxious [SR]	.393	-.029 (.022)	-.061 (.047)	.456	2496	-.008 (.024)	-.013 (.040)	.411	2037
Calm and peaceful [SR]	.466	.061* (.022)	.131* (.047)	.443	2530	.014 (.024)	.024 (.040)	.487	2069
Sleeps at least 7 and <9 hours per night [SR]	.450	.033 (.022)	.070 (.048)	.447	2503	.016 (.025)	.026 (.041)	.443	2046
B. Physical health									
Has fair or poor health [SR]	.330	.017 (.019)	.036 (.041)	.295	2530	.011 (.021)	.019 (.036)	.310	2073
Has trouble carrying groceries or climbing stairs [SR]	.436	-.018 (.021)	-.039 (.045)	.423	2526	-.020 (.023)	-.034 (.038)	.418	2070
Had an asthma or wheezing attack during past year [SR]	.212	-.013 (.018)	-.027 (.038)	.206	2529	-.010 (.019)	-.017 (.032)	.208	2071
Obese, BMI≥30 [SR]	.468	-.048* (.022)	-.103* (.047)	.502	2450	-.046 (.025)	-.077 (.041)	.491	1999
Has hypertension, [M] SBP≥140 or DBP≥90	.297	.022 (.020)	.048 (.045)	.241	2315	.022 (.023)	.037 (.039)	.267	1900

Notes. M = direct measurement. SR = self-report. CM = control mean. SBP = systolic blood pressure. DBP = diastolic blood pressure. Intent-to-treat (ITT) from equation (1), using covariates in Table C1 and weights described in the text. TOT = Treatment-on-treated from equation (2) estimated by two stage least squares with treatment group assignment indicator variables as the instruments for the treatment take-up indicator variables. CCM = control complier mean, as defined in the text. * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses.

Table F6. Effects on Economic Self-sufficiency and Health by Age at Randomization

	Age < 33 at RA			Age ≥ 33 at RA			Diff. by Age	
	CM	E-C ITT	S-C ITT	CM	E-C ITT	S-C ITT	E-C ITT	S-C ITT
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
A. Economic self-sufficiency								
Adult employed and not on TANF [SR]	.467	.050	.021	.439	-.013	.010	-.063	-.011
		(.030)	(.033)		(.028)	(.031)	(.040)	(.045)
Employed [SR]	.555	.032	.032	.484	-.001	.015	-.033	-.017
		(.030)	(.033)		(.028)	(.031)	(.041)	(.045)
Earnings in 2001 [SR]	9643	589	-508	7980	-362	486	-951	994
		(659)	(691)		(609)	(689)	(896)	(980)
Receiving TANF [SR]	.334	-.036	-.042	.254	-.006	-.020	.030	.022
		(.027)	(.030)		(.026)	(.028)	(.037)	(.040)
Income received from government sources during 2001 [SR]	2420	-84	-382	2552	479	179	563	561
		(252)	(277)		(269)	(295)	(370)	(400)
B. Mental health								
Psychological distress, K6 z-score [SR]	-.021	-.090	-.051	.125	-.095	-.012	-.005	.039
		(.064)	(.069)		(.067)	(.075)	(.092)	(.102)
Probability of major depressive episode [SR]	.153	-.021	-.013	.177	-.035	-.014	-.014	-.001
		(.020)	(.021)		(.021)	(.024)	(.029)	(.032)
Worried, tense, or anxious [SR]	.360	-.015	.026	.429	-.043	-.043	-.028	-.069
		(.030)	(.033)		(.031)	(.034)	(.043)	(.047)
Calm and peaceful [SR]	.474	.051	.025	.457	.073*	.003	.022	-.023
		(.031)	(.033)		(.031)	(.035)	(.044)	(.048)
Sleeps at least 7 and <9 hours per night [SR]	.479	.045	.027	.420	.020	.005	-.026	-.021
		(.031)	(.034)		(.032)	(.035)	(.045)	(.049)
C. Physical health								
Has fair or poor health [SR]	.248	-.012	-.030	.416	.046	.054	.057	.084*
		(.026)	(.028)		(.029)	(.033)	(.039)	(.043)
Has trouble carrying groceries or climbing stairs [SR]	.332	-.038	-.043	.545	.001	.002	.039	.045
		(.029)	(.031)		(.030)	(.033)	(.042)	(.045)
Had an asthma or wheezing attack during past year [SR]	.205	-.028	-.031	.221	.003	.011	.031	.042
		(.025)	(.026)		(.025)	(.028)	(.035)	(.038)
Obese, BMI≥30 [SR]	.452	-.056	-.069*	.484	-.040	-.023	.015	.047
		(.031)	(.034)		(.032)	(.035)	(.044)	(.049)
Has hypertension, [M] SBP≥140 or DBP≥90	.227	-.030	-.010	.369	.075*	.055	.104*	.064
		(.027)	(.030)		(.031)	(.035)	(.041)	(.046)

Notes. SR = self-report. M = direct measurement. SBP = systolic blood pressure. DBP = diastolic blood pressure. RA = random assignment. CM = control mean. E-C: Experimental - Control. S-C: Section 8 - Control. Intent-to-treat (ITT) from equation (2), using covariates in Table C1 and weights described in the text, where *X* also contains an indicator for age <33 and *Z* contains interactions of age<33 and age≥33 with the treatment indicator. The total number of completed surveys was 1793 for adults under age 33 and 1733 for those 33 and older. * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses.

Table F7. Effects on Employment by Age at Randomization ─ Administrative Data

	Age < 33 at RA			Age ≥ 33 at RA			Diff. by Age	
	CM	E-C ITT	S-C ITT	CM	E-C ITT	S-C ITT	E-C ITT	S-C ITT
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
A. Fraction of quarters employed by calendar year, 4 states								
1998 [ADMIN]	.473	-.015	-.003	.378	-.022	.030	-.008	.034
		(.024)	(.027)		(.024)	(.027)	(.034)	(.038)
1999 [ADMIN]	.520	.010	-.011	.394	-.006	.050	-.017	.061
		(.024)	(.028)		(.024)	(.028)	(.034)	(.039)
2000 [ADMIN]	.537	.055*	.011	.440	-.009	.026	-.063	.014
		(.025)	(.028)		(.026)	(.028)	(.036)	(.040)
2001 [ADMIN]	.549	.029	.030	.456	-.017	-.002	-.046	-.031
		(.026)	(.029)		(.026)	(.029)	(.036)	(.041)
B. Fraction of quarters employed by year since RA, 4 states [ADMIN]								
Year 1 After RA [ADMIN]	.363	-.036	-.057*	.285	-.007	.016	.029	.072*
		(.022)	(.024)		(.022)	(.024)	(.031)	(.034)
Year 2 After RA [ADMIN]	.433	-.025	-.055	.324	.005	.033	.030	.088*
		(.025)	(.029)		(.025)	(.027)	(.035)	(.039)
Year 3 After RA [ADMIN]	.462	.024	-.019	.375	-.000	.032	-.024	.051
		(.026)	(.030)		(.026)	(.028)	(.037)	(.041)
Year 4 After RA [ADMIN]	.490	.055*	.029	.407	.002	.055	-.052	.026
		(.027)	(.030)		(.026)	(.030)	(.038)	(.042)
Year 5 After RA [ADMIN]	.544	.055*	-.005	.439	-.013	.023	-.068	.028
		(.027)	(.031)		(.027)	(.030)	(.038)	(.043)
C. Employment in 2001 from administrative versus survey data								
Any positive earnings in 2001, 4 states [ADMIN]	.670	.017	.055	.549	-.017	-.020	-.035	-.075
		(.028)	(.031)		(.028)	(.031)	(.039)	(.044)
Any positive earnings in 2001, 4 states [SR]	.690	.007	.006	.570	-.006	.006	-.013	.001
		(.033)	(.039)		(.034)	(.038)	(.047)	(.055)
Any positive earnings in 2001, 5 states [SR]	.701	.013	.014	.578	-.027	-.007	-.040	-.021
		(.028)	(.034)		(.029)	(.033)	(.041)	(.047)

Notes. ADMIN = administrative data. SR = self-report. RA = random assignment. CM = control mean. Intent-to-treat (ITT) from equation (1), using covariates in Table C1 and weights described in the text, where *X* also contains an indicator for age <33 and *Z* contains interactions of age<33 and age≥33 with the treatment indicator. * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses. Administrative data on individual earnings and employment are from California, Illinois, Maryland, and New York unemployment insurance records. Records were obtained for 1615 adults less than 33 years old and 1560 adults 33 and older.

Table F8. Effects on Earnings by Age at Randomization ─ Administrative Data

	Age < 33 at RA			Age ≥ 33 at RA			Diff. by Age	
	CM	E-C ITT	S-C ITT	CM	E-C ITT	S-C ITT	E-C ITT	S-C ITT
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
A. Annual earnings by calendar year, 4 states								
1998 [ADMIN]	5377	68	-220	5140	8	588	-60	808
		(399)	(439)		(454)	(496)	(605)	(659)
1999 [ADMIN]	6596	490	-302	6000	244	958	-246	1260
		(486)	(512)		(539)	(594)	(728)	(781)
2000 [ADMIN]	7630	1117*	-247	6956	-171	253	-1288	501
		(540)	(574)		(547)	(594)	(768)	(822)
2001 [ADMIN]	8870	480	-441	7252	-348	344	-828	785
		(608)	(662)		(555)	(630)	(820)	(909)
B. Annual earnings by year since RA, 4 states								
Year 1 After RA [ADMIN]	3885	-489	-857*	3571	34	330	523	1187*
		(350)	(360)		(376)	(413)	(514)	(543)
Year 2 After RA [ADMIN]	4995	-377	-950*	4581	441	430	818	1380*
		(436)	(439)		(494)	(495)	(665)	(661)
Year 3 After RA [ADMIN]	5692	544	-438	5314	381	850	-163	1288
		(490)	(509)		(518)	(558)	(718)	(757)
Year 4 After RA [ADMIN]	6595	1011	256	6199	-68	1049	-1078	793
		(560)	(585)		(555)	(624)	(791)	(851)
Year 5 After RA [ADMIN]	7727	1748*	300	7276	-538	444	-2285*	144
		(610)	(644)		(594)	(684)	(857)	(929)
C. Earnings in 2001 from administrative versus survey data								
Earnings in 2001, 4 states [ADMIN]	8870	480	-441	7252	-348	344	-828	785
		(608)	(662)		(555)	(630)	(820)	(909)
Earnings in 2001, 4 states [SR]	8869	864	-765	7550	4	515	-861	1280
		(718)	(746)		(675)	(778)	(982)	(1087)
Earnings in 2001, 5 states [SR]	9643	589	-508	7980	-362	486	-951	994
		(659)	(691)		(609)	(689)	(896)	(980)

Notes. ADMIN = administrative data. SR = self-report. RA = random assignment. CM = control mean. E-C: Experimental - Control. S-C: Section 8 - Control. Intent-to-treat (ITT) from equation (1), using covariates in Table 1 and weights described in the text, where X also contains an indicator for age <33 and Z contains interactions of age<33 and age≥33 with the treatment indicator. * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses. Administrative data on individual earnings and employment are from California, Illinois, Maryland, and New York unemployment insurance records. Records were obtained for 1615 adults less than 33 years old and 1560 adults 33 and older.

Table F9. Effects on Voucher Use, Housing and Neighborhood Quality, and Safety

		Experimental - Control		Section 8 - Control	
	CM	ITT	N	ITT	N
	(i)	(ii)	(iii)	(iv)	(v)
A. Used MTO Voucher					
Moved using MTO voucher [ADDR]	.000	.467*	2533	.602*	2073
		(.015)		(.017)	
B. Census tract characteristics					
Average census tract poverty rate [ADDR]	.448	-.119*	2533	-.097*	2073
		(.007)		(.006)	
Average census tract share on public assistance [ADDR]	.228	-.063*	2533	-.055*	2073
		(.004)		(.004)	
Average census tract share of adults employed [ADDR]	.384	.074*	2532	.056*	2072
		(.004)		(.004)	
Average census tract share workers in professional and managerial occupations [ADDR]	.215	.041*	2530	.016*	2071
		(.004)		(.004)	
C. Neighborhood quality					
Very or somewhat satisfied with neighborhood [SR]	.476	.136*	2510	.106*	2056
		(.022)		(.024)	
Neighborhood problems index [SR]	.539	-.126*	2510	-.093*	2056
		(.017)		(.019)	
Negative exterior conditions of buildings and neighborhood [OBS]	.201	-.038*	2359	-.029*	1921
		(.011)		(.012)	
D. Safety					
Streets are safe or very safe during the day [SR]	.758	.090*	2501	.090*	2049
		(.018)		(.018)	
Streets are safe or very safe at night [SR]	.554	.141*	2480	.091*	2031
		(.022)		(.024)	
Member of household victimized by crime during past 6 months [SR]	.213	-.042*	2530	-.055*	2071
		(.017)		(.018)	
Saw drugs sold or used during past 30 days [SR]	.457	-.118*	2481	-.104*	2023
		(.022)		(.024)	
Police not coming when called is a problem in the neighborhood [SR]	.342	-.128*	2338	-.096*	1913
		(.020)		(.023)	
E. Housing quality					
Unit is in poor or fair condition [SR]	.473	-.096*	2504	-.067*	2051
		(.022)		(.024)	
Home problems index [SR]	.340	-.050*	2512	-.027	2059
		(.013)		(.014)	
Interior of the home negative conditions index [OBS]	.190	-.013	2397	-.016	1950
		(.010)		(.011)	
Exterior of the home negative conditions index [OBS]	.170	-.034*	2415	-.028*	1969
		(.011)		(.012)	

Notes. ADDR = address history from tracking file linked to Census data on tract characteristics. Census tract characteristics are averaged across individual addresses since RA, weighted by duration. Except for "managerial and professional occupations" (for which only 2000 Census data was used due to differences in 1990 and 2000 occupation classifications), the characteristics of an address are a linear interpolation from the 1990 Census and 2000 Census. SR = self-report. OBS = interviewer observations. CM = control mean. Intent-to-treat (ITT) from equation (1), using covariates in Table C1 and weights described in the text. * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses.

Table F10. Effects on Social Networks

		Experimental - Control		Section 8 - Control	
	CM	ITT	N	ITT	N
	(i)	(ii)	(iii)	(iv)	(v)
Has three or more close friends [SR]	.351	.017	2525	.006	2071
		(.021)		(.023)	
Visits friends or relatives in their home at least once a week [SR]	.426	-.023	2525	-.021	2064
		(.022)		(.024)	
Visits friends or relatives in own home at least once a week [SR]	.428	-.023	2525	.006	2061
		(.022)		(.024)	
Has diffuse network of friends in which only a few friends know each other [SR]	.276	-.016	2520	.025	2062
		(.019)		(.022)	
Found current job through a friend, relative or acquaintance living in neighborhood [SR]	.075	.002	2490	.018	2041
		(.012)		(.013)	
Has no friends who live in the neighborhood [SR]	.588	.022	2527	.048*	2067
		(.022)		(.024)	
Chats with neighbor at least once a week [SR]	.492	.020	2523	.015	2064
		(.022)		(.024)	
Has a friend who graduated college or earns more than $30,000 a year [SR]	.518	.053*	2334	.032	1917
		(.022)		(.025)	
Attends church or religious service at least once a month [SR]	.426	-.031	2521	.008	2064
		(.021)		(.024)	
Believes people can be trusted [SR]	.097	.011	2505	.006	2056
		(.014)		(.015)	
Experienced discrimination in a shop, restaurant, the neighborhood, child's school, or by police during the past 6 months [SR]	.244	-.038*	2532	-.045*	2072
		(.018)		(.019)	

Notes. SR = self-report. CM = control mean. Intent-to-treat (ITT) from equation (1), using covariates in Table C1 and weights described in the text. * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses.

Table F11. Effects on Education, Training, Health Behaviors and Health Care Access

		Experimental versus Control		Section 8 versus Control	
	CM	ITT	N	ITT	N
	(i)	(ii)	(iii)	(iv)	(v)
A. Education and training					
Years of education completed [SR]	11.047	.020	2516	-.104	2057
		(.096)		(.107)	
Has high school diploma or GED [SR]	.586	-.006	2524	.020	2063
		(.017)		(.020)	
Participated in job training since September 2000 [SR]	.181	-.018	2523	.017	2064
		(.016)		(.019)	
B. Exercise and nutrition					
Moderate physical exercise, fraction of week engaged in [SR]	.471	.025	2516	.049*	2064
		(.018)		(.020)	
Diet, fraction of week ate green vegetables or fruit [SR]	.670	.030*	2511	.019	2059
		(.014)		(.015)	
C. Smoking and drinking					
Smoking [SR]	.293	.010	2512	.005	2059
		(.020)		(.022)	
Binge drinking during past year [SR]	.073	.003	2483	.006	2035
		(.012)		(.013)	
D. Health care access					
Has health insurance [SR]	.849	.018	2528	.006	2067
		(.017)		(.018)	
Has a usual place to go when sick [SR]	.945	-.008	2530	.011	2072
		(.011)		(.011)	

Notes. SR = self-report. CM = control mean. Intent-to-treat (ITT) from equation (1), using covariates in Table C1 and weights described in the text. * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses.

Table F12. Effects on Mobility and Housing Assistance, Access to Transportation, & Relative Income

		Experimental versus Control		Section 8 versus Control	
	CM	ITT	N	ITT	N
	(i)	(ii)	(iii)	(iv)	(v)
A. Mobility and housing assistance					
Moved more than 10 miles from baseline address [ADDR]	.106	.054* (.016)	2424	.028 (.018)	2005
Currently receiving Section 8 [SR]	.255	.294* (.021)	2317	.345* (.023)	1907
Lives in baseline neighborhood or still has friends there [SR]	.746	-.050* (.020)	2526	-.086* (.021)	2065
Lives in baseline neighborhood or has friends from there who come to visit at least a couple of times a year [SR]	.613	-.083* (.022)	2525	-.080* (.023)	2064
Lives in baseline neighborhood or goes back to visit at least a couple of times a year [SR]	.664	-.071* (.021)	2522	-.084* (.022)	2062
B. Access to transportation					
Takes less than 15 minutes to get to nearest bus or train stop [SR]	.921	.015 (.012)	2493	-.003 (.015)	2042
Someone in household has a car, van or truck that runs [SR]	.381	.011 (.020)	2529	.026 (.022)	2070
Has a valid driver's license [SR]	.454	.016 (.020)	2532	-.002 (.023)	2072
C. Relative income					
Household income as fraction of median household income for the tract [SR]	.811	-.177* (.033)	2220	-.174* (.034)	1817

Notes. ADDR = address history from tracking file and linked to Census data. SR = self-report. CM = control mean. Intent-to-treat (ITT) from equation (1), using covariates in Table C1 and weights described in the text. Relative income is household income from 2001 divided by the median household income for the Census tract for the year 1999 in 2001 dollars. * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses.

Table F13. Baseline Characteristics of Adult Survey Respondents and the Full Adult Sample

Variable	Controls		Experimental		Section 8	
	Respon-dents (i)	Full Sample (ii)	Respon-dents (iii)	Full Sample (iv)	Respon-dents (v)	Full Sample (vi)
Demographics						
Age in years (as of December 2001)	39.6	39.6	39.7	39.9	40.1	40.0
Male	.02	.02	.01	.02*	.02	.02
Baltimore site	.15	.15	.15	.15	.15	.15
Boston site	.21	.23	.22	.23	.22	.23
Chicago site	.22	.21	.23	.21*	.23	.21*
Los Angeles site	.16	.16	.16	.16	.15	.16
New York site	.25	.25	.25	.25	.25	.25
African-American	.66	.66	.67	.64*	.66	.64*
Other race	.27	.27	.26	.28*	.26	.27
Hispanic ethnicity, any race	.29	.30	.29	.30*	.30	.31
Never married	.62	.63	.62	.62	.62	.62
Teen parent	.24	.25	.25	.25	.26	.26
Economic and education						
Working	.25	.25	.29	.27	.25	.25
On AFDC	.75	.74	.74	.74	.75	.75
In school	.16	.16	.16	.16	.16	.17
High school diploma	.38	.38	.41	.42	.41	.40
General equivalency diploma	.21	.22	.18	.18	.19	.20
Household						
Had car	.15	.15	.17	.18	.16	.17
Household member with a disability	.16	.16	.16	.16	.17	.15
Household member victimized by crime during past 6 months	.41	.41	.42	.43	.43	.42
No teen children	.62	.63	.59	.60	.61	.61
Household of size 2	.20	.20	.23	.22	.21	.21
Household of size 3	.32	.32	.30	.30	.31	.30
Household of size 4	.22	.22	.23	.23	.23	.23
Neighborhood and housing						
Lived in neighborhood 5+ years	.62	.61	.61	.60*	.63	.62
Moved > 3 times in past 5 years	.11	.11	.08	.09	.09	.09
Very dissatisfied with neighborhood	.46	.47	.46	.47	.47	.47
Streets very unsafe at night	.49	.50	.48	.49	.49	.50
Chats with neighbors 1+ / week	.55	.54	.52	.52	.50	.50
Very likely to tell neighbor if saw their child getting into trouble	.56	.57	.53	.55*	.55	.54
No family living in neighborhood	.65	.65	.65	.65	.62	.63
No friends living in neighborhood	.41	.41	.40	.41	.38	.39
Very sure would find an apartment in another part of city	.45	.45	.45	.46	.48	.49
To get away from gangs or drugs was primary or secondary reason for moving	.78	.78	.77	.77	.75	.76
Better schools was primary or secondary reason for moving	.48	.47	.47	.47	.52	.51
Had applied for S8 voucher before	.45	.44	.41	.42	.39	.39
N	1080	1310	1453	1729	993	1209

Notes: S8 = Section 8. Table consists of the covariates included in the regression models; age is included in the model as a sixth order Legendre polynomial rather than in years. * = p-value < .05 on difference between respondents and full sample.

Table F14. Effects on Change in Employment within Zip Code Between 1994 and 2001

		Experimental versus Control		Section 8 versus Control	
	CM	ITT	N	ITT	N
	(i)	(ii)	(iii)	(iv)	(v)
Residence 1 year after RA					
Change from 1994 to 1995 in log of employment [ADDR]	-.008	.010*	2462	.013*	2028
		(.003)		(.003)	
Change from 1994 to 1996 in log of employment [ADDR]	-.023	.005	2462	-.000	2028
		(.005)		(.006)	
Change from 1994 to 1997 in log of employment [ADDR]	-.028	.015*	2462	-.002	2028
		(.007)		(.007)	
Change from 1994 to 1998 in log of employment [ADDR]	-.011	.007	2462	-.006	2028
		(.007)		(.008)	
Change from 1994 to 1999 in log of employment [ADDR]	.015	.005	2462	-.012	2028
		(.008)		(.009)	
Change from 1994 to 2000 in log of employment [ADDR]	.056	.001	2462	-.029*	2028
		(.009)		(.010)	
Change from 1994 to 2001 in log of employment [ADDR]	.065	.001	2462	-.032*	2028
		(.009)		(.010)	
Residence in 2002					
Change from 1994 to 1995 in log of employment [ADDR]	.005	.004	2453	.012*	2021
		(.003)		(.005)	
Change from 1994 to 1996 in log of employment [ADDR]	-.009	-.006	2453	.005	2021
		(.007)		(.007)	
Change from 1994 to 1997 in log of employment [ADDR]	-.014	.004	2453	.005	2021
		(.008)		(.009)	
Change from 1994 to 1998 in log of employment [ADDR]	.001	.003	2453	.001	2021
		(.009)		(.009)	
Change from 1994 to 1999 in log of employment [ADDR]	.024	.002	2453	-.003	2021
		(.010)		(.010)	
Change from 1994 to 2000 in log of employment [ADDR]	.050	.002	2453	-.007	2021
		(.010)		(.011)	
Change from 1994 to 2001 in log of employment [ADDR]	.050	-.001	2453	-.006	2021
		(.011)		(.011)	

Notes. RA = randomization. ADDR = address history from tracking file linked to zip code-level employment data. Employment data is from the U.S. Census Bureau's Zip Code Business Patterns for 1994 through 2001. Total employment represents the sum of full- and part-time employees on the payroll of establishments in the zip code. For zip codes with suppressed employment data, employment was imputed as the midpoint of the relevant range indicated by the data suppression flag. Change in employment is defined as the log of employment in the later year minus the log of employment in 1994. CM = control mean. Intent-to-treat (ITT) from equation (1), using covariates in Table C1 and weights described in the text. * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses.

G. Additional results for youth

Table G1. List of Tables for Youth Results

Table	Title
G2	Effects for Youth Outcomes
G3	Effects for Neighborhood and Victimization Mediators
G4	Effects for Housing, Parenting, and School Mediators
G5	Effects for Peer and Adult Role Model Mediators
G6	Effects for Educational Mediators
G7	Effects for Health Mediators
G8	Effects for Residential Mobility

Table G2. Intent-To-Treat Effects for Youth Outcomes

Outcome	Female			Male			Male - Female	
	CM (1)	E-C (2)	S-C (3)	CM (4)	E-C (5)	S-C (6)	E-C (7)	S-C (8)
A. Physical health								
Overall health fair/poor [SR]	.101	.008	-.017	.045	.033	.027	.025	.044
		(.029)	(.029)		(.019)	(.025)	(.035)	(.038)
Asthma attack in past year [SR]	.201	.002	-.048	.122	.016	.039	.014	.088
		(.037)	(.038)		(.032)	(.039)	(.049)	(.056)
Non-sport injury in past year [SR]	.115	-.015	-.028	.062	.087*	.080*	.102*	.108*
		(.025)	(.026)		(.026)	(.028)	(.036)	(.039)
Body Mass Index > 95th percentile [SR]	.173	-.009	-.042	.161	.026	-.012	.036	.030
		(.034)	(.037)		(.037)	(.041)	(.049)	(.055)
B. Mental health								
Psychological distress -- K6 scale z-score [SR]	.268	-.289*	-.145	-.162	.095	.005	.385*	.150
		(.094)	(.106)		(.085)	(.100)	(.125)	(.143)
Ever had serious depression symptoms [SR]	.137	-.055	-.061	.031	.013	-.005	.068	.056
		(.030)	(.032)		(.022)	(.024)	(.037)	(.040)
Ever had generalized anxiety symptoms [SR]	.121	-.069*	-.075*	.055	-.015	-.049*	.054	.026
		(.027)	(.029)		(.024)	(.024)	(.036)	(.038)
C. Education								
Graduated HS or still in school [PRY]	.772	.064	.049	.759	-.044	-.040	-.108*	-.090
		(.036)	(.037)		(.037)	(.041)	(.051)	(.055)
In school or working [SR]	.771	.040	-.019	.758	.018	-.007	-.022	.012
		(.035)	(.037)		(.035)	(.040)	(.050)	(.054)
WJ-R reading z-score [WJR]	.059	.093	.046	-.110	-.087	.048	-.180	.002
		(.084)	(.092)		(.096)	(.111)	(.125)	(.142)
WJ-R math z-score [WJR]	.005	.119	.071	-.042	-.095	.019	-.214	-.052
		(.095)	(.099)		(.097)	(.107)	(.132)	(.145)
D. Risky behavior								
Used marijuana in past 30 days [SR]	.131	-.065*	-.072*	.118	.051	.055	.115*	.127*
		(.029)	(.032)		(.030)	(.035)	(.041)	(.047)
Smoked cigarettes in past 30 days [SR]	.191	-.054	-.055	.125	.103*	.151*	.157*	.206*
		(.033)	(.036)		(.032)	(.037)	(.046)	(.052)
Had alcohol in past 30 days [SR]	.206	-.060	-.091*	.140	.063	.061	.122*	.151*
		(.037)	(.038)		(.033)	(.037)	(.049)	(.052)
Ever pregnant or gotten someone pregnant [SR]	.267	-.011	.036	.119	.028	.032	.039	-.004
		(.040)	(.040)		(.031)	(.035)	(.051)	(.052)

Notes. CM: Control mean. E-C: experimental - control difference. S-C: Section 8 - control difference. SR: Self-report. PRY: Parental report about youth. WJR: Woodcock Johnson Revised Assessment. Differences regression-adjusted, with standard errors clustered by household. * indicates p-value <.05. Surveys were completed in experimental, Section 8, and control groups with 749, 510, and 548 respondents respectively ages 15-20 on 12/31/2001.

Table G3. Intent-To-Treat Effects for Neighborhood and Victimization Mediators

Outcome	Female			Male			Male - Female	
	CM (1)	E-C (2)	S-C (3)	CM (4)	E-C (5)	S-C (6)	E-C (7)	S-C (8)
A. General Neighborhood								
Youth lives in baseline neighborhood [SR]	.455	-.143* (.043)	-.148* (.046)	.485	-.101* (.045)	-.120* (.048)	.042 (.059)	.028 (.064)
Poverty rate in current neighborhood [ADDRESS]	.402	-.086* (.017)	-.071* (.016)	.396	-.088* (.019)	-.064* (.018)	-.002 (.025)	.008 (.024)
Pct minority in neighborhood [ADDRESS]	.877	-.033 (.020)	.017 (.017)	.869	-.035 (.021)	-.041 (.022)	-.001 (.028)	-.058* (.027)
Pct youth in neighbrhood not in school or work [ADDRESS]	.120	-.014 (.008)	-.014 (.008)	.126	-.016 (.010)	-.015 (.011)	-.003 (.012)	-.001 (.013)
Pct adults in pro/mgmt occupations [ADDRESS]	.206	.042* (.010)	.016 (.009)	.224	.020 (.011)	.002 (.010)	-.022 (.014)	-.014 (.013)
Not satisfied with neighborhood [PR]	.555	-.177* (.051)	-.203* (.054)	.511	-.100 (.051)	-.059 (.054)	.078 (.073)	.143 (.078)
Feels unsafe in neighborhood at night [PR]	.437	-.170* (.047)	-.086 (.054)	.509	-.178* (.049)	-.155* (.052)	-.008 (.070)	-.069 (.076)
Fraction of 4 types of discrimination in ‘hood [SR]	.107	-.018 (.019)	-.006 (.019)	.134	-.024 (.019)	.005 (.025)	-.006 (.027)	.011 (.030)
Fraction of 6 problems with neighborhood [PR]	.565	-.164* (.037)	-.128* (.041)	.509	-.102* (.039)	-.064 (.037)	.062 (.054)	.064 (.056)
Saw drugs in neighborhood 1+/week in past 30 days [SR]	.437	-.104* (.047)	-.122* (.051)	.441	-.042 (.047)	-.026 (.056)	.062 (.066)	.095 (.075)
Heard gunshots in 'hood 1+/week in past 30 days [SR]	.118	-.040 (.031)	-.053 (.028)	.155	-.034 (.031)	-.075* (.032)	.006 (.043)	-.022 (.043)
B. Victimization								
Household member was crime victim past 6 mths [PR]	.275	-.082 (.044)	-.096* (.043)	.247	-.014 (.044)	-.058 (.046)	.068 (.062)	.038 (.063)
Saw someone shot or stabbed in past 12 mths [SR]	.150	-.043 (.036)	-.047 (.034)	.209	-.016 (.039)	-.030 (.046)	.027 (.053)	.017 (.057)
Was “jumped” in past 12 months [SR]	.085	-.006 (.029)	.005 (.028)	.181	.010 (.038)	-.003 (.041)	.015 (.047)	-.008 (.049)

Notes. CM: Control mean. E-C: experimental - control difference. S-C: Section 8 - control difference. ADDRESS: Address history from tracking file, linked to Census. SR: Self-report. PRY: Parental report about youth. PR: Parental report about household. Differences regression-adjusted, with standard errors clustered by household. * indicates p-value <.05. For PR measures, analysis was conducted at household level using household average right-hand side variables. Surveys were completed in experimental, Section 8, and control groups with 749, 510, and 548 respondents respectively ages 15-20 on 12/31/2001. Types of discrimination were: at school or work, neighborhood recreation program, shopping or restaurant, with police. Problems with neighborhood were: litter, graffiti, public drinking, abandoned buildings, people hanging out, police not coming. Types of criminal victimization were: purse or wallet snatched, threatened with weapon, beaten or assaulted, break-in to home, stabbed or shot.

Table G4. Intent-To-Treat Effects for Housing, Parenting, and School Mediators

Outcome	Female			Male			Male - Female	
	CM (1)	E-C (2)	S-C (3)	CM (4)	E-C (5)	S-C (6)	E-C (7)	S-C (8)
A. Housing								
Overall housing condition is fair/poor [PR]	.477	-.071 (.049)	-.020 (.054)	.507	-.055 (.050)	-.098 (.053)	.016 (.071)	-.078 (.076)
Fraction of 7 problems with home [PR]	.334	-.048 (.025)	-.035 (.029)	.333	-.052 (.028)	-.038 (.029)	-.004 (.038)	-.003 (.042)
Fraction of 7 problems with home interior [OBS]	.216	-.055* (.022)	-.013 (.026)	.222	-.022 (.025)	-.030 (.030)	.033 (.035)	-.017 (.040)
Fraction of 7 problems with home exterior [OBS]	.218	-.062* (.024)	-.039 (.029)	.227	-.037 (.026)	-.034 (.029)	.024 (.037)	.006 (.039)
B. Parenting Practices								
Mother /primary caregiver is very supportive [SR]	.670	.035 (.045)	.024 (.049)	.842	-.056 (.034)	-.054 (.039)	-.091 (.056)	-.078 (.063)
Parent knows all about friends & whereabouts [SR]	.258	.012 (.039)	-.050 (.041)	.173	-.044 (.036)	-.034 (.041)	-.056 (.054)	.016 (.059)
No adult present after school [SR]	.242	.046 (.050)	-.010 (.055)	.301	.068 (.051)	.061 (.059)	.022 (.070)	.071 (.079)
Fraction days/week family eats together [PR]	.571	.042 (.039)	.034 (.040)	.596	-.021 (.040)	.016 (.042)	-.063 (.058)	-.019 (.059)
Fraction of 4 types of parental contact w/schl [PR]	.370	.022 (.031)	.023 (.033)	.418	-.034 (.033)	-.010 (.033)	-.056 (.046)	-.033 (.047)
C. School Environment								
% free lunch [ADMIN]	.516	-.053* (.022)	.000 (.022)	.524	-.079* (.024)	-.033 (.026)	-.026 (.032)	-.033 (.033)
% limited English proficient [ADMIN]	.155	-.030* (.014)	-.004 (.016)	.163	-.033* (.014)	-.032* (.015)	-.003 (.019)	-.028 (.021)
% white [ADMIN]	.114	.061* (.020)	.012 (.020)	.112	.065* (.021)	.067* (.026)	.004 (.029)	.055 (.033)
Pupil-teacher ratio [ADMIN]	18.6	.533 (.358)	-.252 (.459)	17.4	1.402* (.414)	.607 (.441)	.868 (.544)	.859 (.633)
Percentile rank on state exam [ADMIN]	.240	.040 (.024)	-.013 (.024)	.188	.063* (.026)	.037 (.027)	.023 (.034)	.050 (.035)
Fraction of 5 positive school climate items [SR]	.621	-.001 (.033)	.039 (.034)	.599	.028 (.031)	-.008 (.037)	.029 (.045)	-.047 (.049)

Notes. CM: Control mean. E-C: experimental - control difference. S-C: Section 8 - control difference. ADMIN: State data on schools. OBS: interviewer observation of housing unit. SR: Self-report. PRY: Parental report about youth. PR: Parental report about household. Differences regression-adjusted, with standard errors clustered by household. * indicates p-value <.05. For PR measures, analysis was conducted at household level using household average right-hand side variables. Surveys were completed in experimental, Section 8, and control groups with 749, 510, and 548 respondents respectively ages 15-20 on 12/31/2001. Problems with home were: peeling paint, plumbing, rats or mice, cockroaches, broken locks, broken windows, heat. Interviewer observations of problems with home interior were: cracks in walls, peeling paint, mold, cigarette smoke, noisy inside, noisy outside, cluttered. Interviewer observations of problems with home exterior were: condition of unit, condition of other units on block, metal bars on unit, metal bars on other units, condition of block, broken windows, junk on block. Items parental knows everything about were: who friends are, who with when not home.

Table G5. Intent-To-Treat Effects for Peer and Adult Role Model Mediators

Outcome	Female			Male			Male - Female	
	CM (1)	E-C (2)	S-C (3)	CM (4)	E-C (5)	S-C (6)	E-C (7)	S-C (8)
A. Peers								
Has at least one close friend [PRY]	.890	.036	-.013	.917	.008	.051	-.028	.064
		(.025)	(.034)		(.027)	(.029)	(.036)	(.044)
Has 5 or more friends [SR]	.382	.050	-.004	.530	.024	.067	-.026	.071
		(.044)	(.046)		(.047)	(.051)	(.064)	(.070)
Friends involved in school activities [SR]	.615	.071	.050	.710	-.013	.010	-.083	-.040
		(.045)	(.048)		(.042)	(.049)	(.062)	(.071)
Has friends who use drugs [SR]	.295	.002	-.001	.327	.127*	.161*	.125*	.161*
		(.043)	(.044)		(.046)	(.051)	(.062)	(.067)
Has friends who carry weapons [SR]	.098	.009	.031	.157	.037	-.033	.028	-.064
		(.026)	(.030)		(.039)	(.036)	(.045)	(.046)
Has relatives or friends who belong to a gang [SR]	.154	.005	-.029	.187	-.063	-.055	-.068	-.027
		(.035)	(.031)		(.032)	(.037)	(.046)	(.047)
Friends from baseline visit new neighborhood [SR]	.178	-.018	-.007	.164	-.022	-.047	-.004	-.040
		(.036)	(.041)		(.035)	(.036)	(.050)	(.055)
Visits baseline 'hood but doesn't live there [SR]	.234	-.026	-.012	.205	.034	.022	.061	.034
		(.037)	(.043)		(.040)	(.044)	(.054)	(.061)
B. Adult Role Models								
Likely neighbors intervene vs. graffiti [PR]	.497	.166*	.105	.575	.052	-.026	-.114	-.131
		(.049)	(.057)		(.050)	(.056)	(.071)	(.080)
Likely neighbors intervene if kids skipping school [PR]	.343	.099	.038	.370	.086	.067	-.013	.029
		(.053)	(.058)		(.050)	(.056)	(.075)	(.083)
Structured activity after school [SR]	.275	.050	.004	.248	.064	.051	.015	.047
		(.043)	(.042)		(.041)	(.047)	(.060)	(.064)
Attended 1+ church youth activities per month [SR]	.380	.006	-.050	.313	-.020	.012	-.025	.062
		(.048)	(.048)		(.041)	(.045)	(.062)	(.065)
Saw father at least once a week in past 12 months [SR]	.253	.068	.078	.365	-.043	-.017	-.111	-.095
		(.040)	(.046)		(.044)	(.047)	(.057)	(.065)
Father has been very supportive [SR]	.235	.026	-.000	.271	.033	-.006	.007	-.006
		(.039)	(.043)		(.041)	(.043)	(.055)	(.060)
Comfortable talking about problems w/3+ adults [SR]	.305	.133*	.061	.397	-.005	-.009	-.138*	-.069
		(.042)	(.048)		(.046)	(.051)	(.062)	(.070)
Has 4+ adults who care and will help if trouble [SR]	.448	.070	.028	.498	.003	-.042	-.067	-.070
		(.047)	(.047)		(.046)	(.053)	(.066)	(.072)

Notes. CM: Control mean. E-C: experimental - control difference. S-C: Section 8 - control difference. SR: Self-report. PRY: Parental report about youth. PR: Parental report about household. Differences regression-adjusted, with standard errors clustered by household. * indicates p-value <.05. For PR measures, analysis was conducted at household level using household average right-hand side variables. Surveys were completed in experimental, Section 8, and control groups with 749, 510, and 548 respondents respectively ages 15-20 on 12/31/2001. No adult present was: no adult at either 3:45, 5:30, or 7:30 on selected day of week. Parental contact with school (for any child in household) was: went to general school meeting, went to a school event, volunteered at school, volunteered for team or club.

Table G6. Intent-To-Treat Effects for Educational Mediators

Outcome	Female			Male			Male - Female	
	CM (1)	E-C (2)	S-C (3)	CM (4)	E-C (5)	S-C (6)	E-C (7)	S-C (8)
A. School Engagement								
Always pays attention in class [SR]	.490	.118* (.056)	-.002 (.062)	.484	-.013 (.056)	.066 (.062)	-.131 (.078)	.068 (.087)
Works hard in school [SR]	.508	.058 (.058)	.028 (.061)	.449	-.101 (.056)	-.013 (.068)	-.158* (.079)	-.040 (.091)
B grades or higher last year [SR]	.415	-.008 (.047)	-.018 (.050)	.293	-.055 (.042)	-.105* (.044)	-.047 (.062)	-.088 (.067)
Always finishes homework [SR]	.505	.027 (.060)	-.003 (.066)	.406	-.029 (.055)	-.078 (.062)	-.057 (.082)	-.074 (.091)
At least 5 hours/week of homework [SR]	.488	.052 (.057)	.045 (.063)	.354	.056 (.053)	.110 (.061)	.005 (.078)	.065 (.087)
At least 5 hours/week of reading [SR]	.377	-.006 (.045)	-.026 (.051)	.250	.023 (.042)	.028 (.049)	.030 (.060)	.054 (.068)
B. Educational Track								
Ever took SAT, ACT, or AP exams [SR]	.426	-.047 (.046)	.037 (.052)	.358	-.037 (.048)	.045 (.053)	.010 (.067)	.009 (.074)
Ever took algebra or higher math [SR]	.833	.012 (.033)	-.005 (.039)	.827	-.085* (.037)	-.055 (.035)	-.097 (.050)	-.050 (.053)
Gifted class in past 2 years [PRY]	.068	.057 (.031)	.010 (.034)	.147	-.040 (.039)	-.039 (.037)	-.097 (.050)	-.049 (.051)
Special education in past 2 years [PRY]	.154	.037 (.038)	-.002 (.038)	.324	.014 (.050)	-.038 (.052)	-.023 (.063)	-.035 (.065)
C. Educational Problems								
Ever repeated a grade [PRY]	.200	.096* (.036)	.009 (.037)	.326	-.028 (.041)	-.049 (.042)	-.124* (.054)	-.058 (.056)
Late for school once a month or more [SR]	.679	-.020 (.042)	.033 (.048)	.616	.030 (.046)	.060 (.051)	.050 (.061)	.027 (.070)
Absent from school 5% or more of the school year [SR]	.426	-.076 (.047)	-.040 (.052)	.389	-.002 (.045)	.005 (.052)	.074 (.065)	.045 (.073)
School requested meet about prob past 2 yrs [PRY]	.184	-.034 (.039)	-.012 (.046)	.337	.044 (.050)	.061 (.058)	.078 (.064)	.072 (.073)
Was suspended/expelled from school past 2 yrs [PRY]	.117	.011 (.035)	.042 (.042)	.301	-.033 (.050)	-.080 (.056)	-.044 (.060)	-.122 (.070)
D. Future Expectations								
Believes chances high will complete college [SR]	.543	.073 (.046)	.034 (.049)	.449	-.044 (.044)	-.053 (.048)	-.117 (.064)	-.088 (.069)
Believes chances high will find good job as adult [SR]	.742	.055 (.037)	.007 (.044)	.652	-.003 (.043)	.039 (.047)	-.059 (.056)	.032 (.065)

Notes. CM: Control mean. E-C: experimental - control difference. S-C: Section 8 - control difference. SR: Self-report. PRY: Parental report about youth. Differences regression-adjusted, with standard errors clustered by household. * indicates p-value <.05. Surveys were completed in experimental, Section 8, and control groups with 749, 510, and 548 respondents respectively ages 15-20 on 12/31/2001. Some items not asked for youth ages 19-20, resulting in smaller sample sizes: no adult present after school, fraction of school climate items, pays attention, works hard, finishes homework, 5+ hours homework, gifted class, special education, school requested meeting, suspended/expelled. School climate was: teachers interested in students, students disruptive, cheating on tests, discipline fair, felt safe. Structured activity was: at school, church or community center -- participating in a sport, club, tutoring, or other organized activity.

Table G7. Intent-To-Treat Effects for Health Mediators

Outcome	Female			Male			Male - Female	
	CM (1)	E-C (2)	S-C (3)	CM (4)	E-C (5)	S-C (6)	E-C (7)	S-C (8)
A. Healthy environment								
Fraction of past 7 days did aerobic exercise [SR]	.353	.055 (.032)	.022 (.040)	.555	.029 (.033)	.020 (.036)	-.027 (.045)	-.003 (.053)
Fraction of past week moderate activity [SR]	.412	.045 (.033)	-.023 (.040)	.476	.067 (.034)	.045 (.037)	.023 (.047)	.068 (.054)
Participates in sport after school [SR]	.032	.061* (.022)	.006 (.017)	.138	.014 (.030)	.033 (.038)	-.047 (.037)	.027 (.042)
Fraction of past 7 days some fruits/vegetables [SR]	.574	.028 (.032)	-.036 (.035)	.568	.010 (.030)	.036 (.032)	-.019 (.044)	.072 (.048)
Fraction of 6 asthma triggers [PR]	.187	.011 (.023)	.021 (.024)	.212	.003 (.021)	.014 (.023)	-.007 (.029)	-.007 (.032)
B. Access to care								
Youth has health insurance [PRY]	.873	-.011 (.029)	.003 (.034)	.809	.081* (.034)	.047 (.039)	.092* (.044)	.044 (.049)
Talked to a doctor about health in past 6 mths [PRY]	.736	.052 (.046)	-.035 (.059)	.728	-.058 (.056)	-.142* (.067)	-.110 (.073)	-.108 (.089)
C. Adult mental health								
Adult distress K6 z-score [PR]	.170	-.275* (.112)	-.193 (.124)	-.058	.160 (.103)	.096 (.110)	.436* (.152)	.289 (.169)
Adult probability of depression [PR]	.187	-.039 (.034)	-.066 (.036)	.138	.003 (.030)	.038 (.037)	.043 (.047)	.104* (.052)
Adult fraction worried, tense or anxious [PR]	.424	-.071 (.050)	-.073 (.055)	.453	-.053 (.051)	-.024 (.056)	.018 (.073)	.049 (.079)
Adult fraction calm and peaceful [PR]	.375	.140* (.049)	.142* (.053)	.508	-.062 (.050)	-.113 (.058)	-.202* (.071)	-.256* (.079)
Adult fraction sleeping 7-8 hours/night [PR]	.362	.118* (.052)	.078 (.056)	.409	.089 (.049)	.058 (.057)	-.028 (.074)	-.020 (.081)

Notes. CM: Control mean. E-C: experimental - control difference. S-C: Section 8 - control difference. SR: Self-report. PRY: Parental report about youth. PR: Parental report about household. Differences regression-adjusted, with standard errors clustered by household. * indicates p-value <.05. For PR measures, analysis was conducted at household level using household average right-hand side variables. Surveys were completed in experimental, Section 8, and control groups with 749, 510, and 548 respondents respectively ages 15-20 on 12/31/2001. Asthma triggers were: rats or mice, cockroaches, wall-to-wall carpet, pets with fur, cigarette smoke, mold.

Table G8. Intent-To-Treat Effects for Residential Mobility

Outcome	Female			Male			Male - Female	
	CM (1)	E-C (2)	S-C (3)	CM (4)	E-C (5)	S-C (6)	E-C (7)	S-C (8)
Program move [ADDRESS]	.000	.471*	.556*	.000	.415*	.568*	-.056	.012
		(.034)	(.039)		(.034)	(.039)	(.047)	(.056)
One or more moves [ADDRESS]	.643	.143*	.203*	.687	.084*	.092*	-.059	-.112*
		(.043)	(.039)		(.036)	(.038)	(.055)	(.052)
One or more moves [PR]	.623	.075	.093*	.654	.012	-.009	-.063	-.103
		(.045)	(.046)		(.047)	(.046)	(.066)	(.065)
Two or more moves [ADDRESS]	.298	.125*	.141*	.301	.079	.159*	-.046	.017
		(.043)	(.047)		(.042)	(.049)	(.058)	(.066)
Two or more moves [PR]	.266	.035	.094	.301	-.023	.063	-.059	-.031
		(.045)	(.050)		(.043)	(.049)	(.065)	(.072)
Number of moves [ADDRESS]	1.04	.374*	.442*	1.14	.208*	.375*	-.166	-.067
		(.091)	(.094)		(.089)	(.105)	(.124)	(.138)
Number of moves [PR]	1.11	.057	.213	1.25	-.101	-.022	-.158	-.234
		(.114)	(.123)		(.120)	(.127)	(.175)	(.173)

Notes to Tables A3-A8. CM: Control mean. E-C: experimental - control difference. S-C: Section 8 - control difference. PR: parental report about household. ADDRESS: Address history from tracking file, linked to Census. ADMIN: Administrative data from school reported to attend or last attended. Differences regression-adjusted, using equation (1) with standard errors clustered by household. * indicates p-value <.05. For PR measures, analysis was conducted at household level using household average right-hand side variables. Sample is ages 15-20 as of 12/31/01.