MOVING TO OPPORTUNITY AND TRANQUILITY:
NEIGHBORHOOD EFFECTS ON ADULT ECONOMIC SELF-SUFFICIENCY AND HEALTH FROM A RANDOMIZED HOUSING VOUCHER EXPERIMENT

Jeffrey R. Kling, Jeffrey B. Liebman, Lawrence F. Katz, and Lisa Sanbonmatsu *

October 2004

ABSTRACT

We study adult economic and health outcomes in the Moving to Opportunity (MTO) demonstration, a randomized housing mobility experiment in which families living in high-poverty U.S. public housing projects in five cities were given vouchers to help them move to private housing units in lower-poverty neighborhoods. An "experimental" group was offered vouchers valid only in a low-poverty neighborhood; a "Section 8" group was offered traditional housing vouchers without geographic restriction; a control group was not offered vouchers. Our sample consists largely of black and Hispanic female household heads with children.

Five years after random assignment, the families offered housing vouchers through MTO lived in safer neighborhoods that had significantly lower poverty rates than those of the control group not offered vouchers. We find no significant overall effects on adult employment, earnings, or public assistance receipt. In contrast, we do find significant mental health benefits of the MTO intervention for the experimental group. We also demonstrate a more general pattern for the mental health results using both voucher groups of systematically larger effect sizes for groups experiencing larger changes in neighborhood poverty rates. In our analysis of physical health outcomes, we find a significant reduction in obesity for the experimental group, but no significant effects on four other aspects of physical health (general health, asthma, physical limitations, and hypertension) or on our summary measure of physical health.

JEL classifications: I18, I38, J38.   Keywords: neighborhood effects; social experiment

This paper studies adult economic and health outcomes in the Moving to Opportunity (MTO) demonstration, a randomized housing mobility experiment in which families living in high-poverty U.S. public housing projects were given vouchers to help them move to private housing units in lower-poverty neighborhoods. Our analysis addresses questions about both the effects of neighborhoods on individual outcomes and the impacts of alternative forms of government housing assistance.

The idea that residence in a distressed community can limit an individual's economic prospects has been powerfully advanced by Wilson (1987, 1996).[1] The location of one's residence may also affect the process of finding employment through the structure of local social interactions (Granovetter 1974; Montgomery 1991, 1992; Topa 2001) and through the impacts of economic or racial segregation (Benabou 1993, Cutler and Glaeser 1997).[2] A wide range of links between residential location and health outcomes have also been hypothesized (Kawachi and Berkman 2003). Of particular relevance to this study, neighborhood slums have long been thought to breed public health problems. Some evidence suggests that the construction of new public housing in the 1950s may have improved health by enabling people to move out of substandard and overcrowded housing conditions (Wilner et al., 1962). During the past quarter century, however, a number of high-poverty, urban public-housing projects have become centers of violent crime and drug use, and concerns have intensified about the health impacts for urban residents of stress, poor air quality, exposure to allergens, and unhealthy diet and exercise habits.

---

[1] Two recent quasi-experimental studies have examined the effects of housing and neighborhood quality on children. Oreopoulos (2003) uses variation in location caused by a local housing authority in Canada to study the impact of children's neighborhoods on economic outcomes later in life. Jacob (2004) uses public housing demolitions in Chicago to study educational outcomes. Also, Sanbonmatsu et al. (2004) uses data from the MTO experiment to examine effects of moving to lower poverty neighborhoods on academic achievement of children. All three find little in the way of neighborhood effects.

[2] Brock and Durlauf (2001) provide an overview of interactions-based models. Ioannides and Loury (2003) survey the role of job information networks.

Non-experimental research has found strong associations between living in disadvantaged neighborhoods and adverse outcomes in both the employment and health domains.[3] Whether these associations reflect a causal relationship between living in disadvantaged neighborhoods and outcomes remains uncertain. People living in high-poverty neighborhoods differ in many ways from those living in lower-poverty neighborhoods, and only some of those differences can be adequately measured and controlled for in non-experimental studies. The random assignment design of MTO encourages otherwise identical people to live in different neighborhoods and thereby provides a rich laboratory for exploring the causal effects of residential location.

This research also bears on whether it is better for the government to provide housing assistance in the form of government operated public housing projects or through vouchers that subsidize rents in the private market. The United States spends about $32 billion per year on housing assistance -- more than on food stamps or cash welfare.[4] Housing projects provide subsidized units for 2.4 million households, and Section 8 vouchers help another 1.6 million households rent private units.[5] Housing assistance is rationed, so that only a minority of eligible families receives assistance and long wait lists for both public housing and vouchers exist in many communities. While the costs of providing a housing project unit and a voucher are similar and require the same contribution of the household toward rent, there could be large

[3] Examples include Kasarda 1989; Jargowsky 1997; Robert 1998; Yen and Kaplan 1999a; Waitzman and Smith 1998; Diez-Roux et al., 2001; Ellaway, Anderson, and Macintyre 1997; Yen and Kaplan 1999b; Ross and Mirowsky 2001; and Silver, Mulvey and Swanson 2002.
[4] Committee on Ways and Means (2004). Cash welfare is combined federal and state spending on Temporary Assistance for Needy Families (TANF) of $24 billion. Food stamps spending is $25 billion. The other major means-tested programs in the U.S. are Medicaid ($258 billion), Supplemental Security Income ($35 billion), and the Earned Income Tax Credit ($34 billion). Numbers are for fiscal year 2002.
[5] Figures are from 1999, with housing projects referring to public housing and project-based Section 8 assistance (Millennial Housing Commission 2002). As housing projects have been demolished and project-based Section 8 contracts terminated, the number of housing vouchers issued has risen. For FY2004, HUD estimates that 2 million vouchers will be funded (U.S. Department of Housing and Urban Development, 2004a).

distortions from providing project-based units instead of vouchers.[6] To receive project-based

assistance, a household must accept a specific bundle of housing and neighborhood attributes. In

contrast, a voucher user can select from a variety of apartments on the private market -- offering

many more choices within the same budget constraint. In particular, public housing projects for

poor families with children in large U.S. metropolitan areas tend to be located in areas with

highly concentrated poverty, while vouchers offer the possibility of living in mixed-income

neighborhoods. If neighborhood attributes affect individual outcomes, then switching from

project-based assistance to vouchers could improve the well-being of those receiving assistance.

The randomized design of MTO provides a direct comparison of outcomes under the two main

forms of U.S. housing assistance.

MTO is a U.S. Department of Housing and Urban Development (HUD) demonstration

conducted in five cities: Baltimore, Boston, Chicago, Los Angeles, and New York. Public

housing residents with children were eligible to participate in a lottery that resulted in random

assignment to one of three groups:[7]

- *Control group*. Received no new assistance, but continued to be eligible for public
  housing.
- *Section 8 group*. Received traditional Section 8 voucher, without geographic restriction.
- *Experimental group*. Received Section 8 voucher, restricted for one year to a census tract
  with a poverty rate of less than 10 percent. Also received mobility counseling.[8]

---

[6] Most studies find that it is somewhat cheaper to provide assistance in the form of vouchers, though there is
considerable disagreement over the magnitude of the cost difference (U.S. Department of Housing and Urban
Development, 2000; Olsen 2000). Tenants in public housing and those using housing vouchers both pay
approximately 30 percent of their income in rent. 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. See Quigley (2000) and Olsen
(2003) for descriptions of alternative forms of housing assistance.

[7] Local housing authorities selected eligible developments from among those in census tracts with poverty rates of
40 percent or higher. Nationally half of all public housing households are elderly or without children, and about one
tenth of all public housing households have children and are in high-poverty census tracts. Approximately one-
quarter of eligible households at MTO sites participated in the lottery (Goering 2003).

[8] The counseling was designed to make it easier for families to locate housing in unfamiliar neighborhoods by
helping families set goals, locate and visit suitable units, negotiate rents, and manage the credit review process.

Random assignment to groups began in 1994. Our sample consists of one adult in each of the 4248 households that were randomly assigned through 1997 at the five sites. 85 percent of these adults are African-American or Hispanic female heads of household with children.[9]

Our analysis draws upon the MTO baseline surveys, state administrative earnings and welfare data, and the 2002 interim evaluation survey of MTO households, which collected data on households four to seven years after enrollment. Earlier research on MTO consists of pilot studies of the participants at individual MTO sites observed within the first few years after random assignment.[10] Relative to these preliminary studies, we have standardized the data collection across all five MTO sites, resulting in a much larger sample and greater external validity for the relevance of results across a variety of settings. We have studied adult physical health for the first time, and have also examined more comprehensive measures of mental health and economic self-sufficiency as well as a broader range of mediating factors to potentially illuminate the mechanisms through which residential neighborhoods may affect economic and health outcomes.

We find that moving out of public housing into neighborhoods with lower poverty rates had no statistically significant effect on economic self-sufficiency, but did significantly improve mental health. In our analysis of physical health outcomes, we find a significant reduction in obesity, but no significant effects on four other aspects of physical health (general health, physical limitations, asthma, and hypertension). There is a non-negligible probability of observing one significant physical health result out of five even if the null hypothesis of no effect

were true, and, using the entire sample, our summary measure of physical health was not significantly affected by the MTO intervention. In our analysis of subgroups, we find a positive and significant impact of the MTO experimental treatment for younger adults, using a summary measure that combines all of the economic, physical health, and mental health outcomes in our study; there was no significant overall impact for older adults.

The remainder of the paper has four sections. Section I describes our analytical methods and data. Section II presents characteristics of our sample, their mobility patterns, and the neighborhoods to which they were exposed. The main results are given in section III. Section IV concludes.

I. Methods

We compare groups of individuals who were similar at the time of random assignment, but who, due to the results of the housing voucher lottery, came to live in different types of neighborhoods. In this way, we use the random assignment of the voucher offer to overcome the standard problem of endogenously determined locations.

A. Regression models

Separate regression models are used to compare the experimental group to the control group and the Section 8 group to the control group. Let Z be an indicator for assignment to a treatment group and D be an indicator for use of a voucher offered through the MTO program. In equation (1), the coefficient π_1 gives the probability of using an offered voucher.

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

Individual characteristics at the time of random assignment are denoted by X.[11]

Let Y_k be the kth individual outcome, such as earnings in the past year. The effect of being offered a housing voucher on an outcome, known as the intent-to-treat (ITT) effect, is the coefficient π_{2k} estimated using equation (2).

(2) $\quad Y_k = Z\pi_{2k} + X\beta_{2k} + \varepsilon_{2k}$

Due to random assignment, Z is orthogonal to X, and the consistency of π_{2k} does not depend upon the inclusion of X in the model. Regression adjustment in equation (2) is used to improve estimation precision and to account for chance differences between groups in the distribution of pre-random assignment characteristics.

As the average effect of an attempted policy intervention on the entire target population, the ITT estimate is often of direct interest. Another parameter of interest is the effect on those who moved using an MTO voucher, known as the effect of the treatment-on-treated (TOT). 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 D in two stage least squares estimation of (3).

(3) $\quad Y_k = D\gamma_{3k} + X\beta_{3k} + \varepsilon_{3k}$

The TOT estimate γ_{3k} is equal to π_{2k}/π_1.[12] 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

---

[11] D is zero for all members of the control group. Equation (1) serves two purposes in this paper. First, models that include one treatment group and the control group are used to estimate π_1 in order to serve as the first stage of the 2SLS model in equation (3) below. In this first set of models, β_1 serves only to adjust for chance differences in X between treatment and control groups and to reduce residual variation. Second, models for factors affecting take-up of the offered voucher are estimated separately for each treatment group (with no control group data). In this second set of models, β_1 is of direct interest and π_1 is not estimated.

[12] Inflation by the proportion in the treatment group who actually received the treatment was introduced in the program evaluation literature and Bloom (1984); an earlier paper on a related topic is by Mallar (1979). See Heckman, LaLonde, and Smith (1999) for a comprehensive discussion of alternative parameters of interest in the evaluation of social programs.

voucher, which we believe is a reasonable approximation, but not strictly true.[13] The fact that

some participants offered vouchers did not use the voucher that was offered also decreases our

statistical power to detect effects on outcomes.

Using terminology of Angrist, Imbens, and Rubin (1996), we refer to those using an

MTO voucher to move as treatment compliers. The treatment complier mean outcome is directly

identified by the data. The implied mean outcome for those in the control group who would have

accepted the treatment if it had been offered to them -- the control complier mean (CCM) -- is

identified under the assumptions needed to estimate the TOT effect (Katz, Kling, and Liebman

2001). CCM is defined in equation (4) for each outcome k.

(4) $\quad CCM_k = E[Y_k \mid D = 1] - \gamma_{3k}$

CCM is the difference between the treatment complier mean and the TOT effect. It is a base rate

for assessing relative changes between treatment and control compliers.

B. *Summary measures*

To form judgments about the overall impact of the intervention on each of our three

outcome domains (economic self-sufficiency, mental health, physical health), we construct

summary measures.[14] The building blocks of the summary measures are standardized treatment

---

[13] 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 those who did not make a subsidized move through the program 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. In this sense, we interpret TOT as a useful approximation. The assumptions for TOT also require that control group members were not affected by the experience of "losing the lottery." We similarly view this as not literally true but as a reasonable approximation. In qualitative research with MTO families, we found that the control group adults have been "lottery losers" numerous times in life and that some do not even recall enrolling in the demonstration.
[14] 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).

effect sizes. Let σ_k^2 equal the variance of Y_k for the control group.[15] Equation (5) defines the

mean effect size, τ, for a set of K outcomes based on the treatment effect estimates and the

control group standard deviations.[16]

$$(5) \qquad \tau = \frac{1}{K}\sum_{k=1}^{K}\frac{\pi_{2k}}{\sigma_k}$$

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 (6). Point estimates for each outcome are identical to those obtained using equation (2).

Let I_K be a K by K identity matrix and let Z and X be defined as in (2).

$$(6) \qquad Y = \left(I_K \otimes \left(Z \quad X\right)\right)\theta + \upsilon \qquad\qquad Y = (Y_1',\ldots,Y_K')'$$

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 (6).[17] A principal reason for reporting summary measures is

to focus on a small number of pre-specified statistical estimates for which we can assess the joint

probability under the null hypothesis of no true effects of observing at least one effect as large as

those observed. We present our methodology for inference in the presence of multiple testing in

conjunction with the results in section III.E.

---

[15] 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.

[16] If there were no missing data on survey items and X contained only a constant, then this summary measure would be identical to using the outcome $Y* = \frac{1}{K}\sum_{k=1}^{K}\frac{Y_k}{\sigma_k}$ and calculating $\tau = E[Y*|Z=1] - E[Y*|Z=0]$. Equation 5 is a more direct summary of the treatment effects from equation (2) in that it incorporates regression adjustment for each outcome as well as data from individuals who have some item nonresponse.

[17] 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.

C. *Data*

The data in this study come from baseline surveys administered prior to random assignment, from administrative data, and from the MTO interim evaluation survey conducted in 2002. The administrative data on earnings and welfare benefits are from the Unemployment Insurance (UI) and state and county welfare agencies of California, Illinois, Maryland, Massachusetts, and New York.[18] The analysis of administrative data reported here was conducted in collaboration with Abt Associates in preparation for our report (Orr et al., 2003) to the U.S. Department of Housing and Urban Development (HUD). The survey design and data collection were also conducted jointly with HUD and Abt Associates. For earnings, data covering 2001 were obtained from both UI records and surveys, and the results were highly but not perfectly correlated.[19] The survey had an effective response rate of 90 percent.[20] Response rates were 91 percent for the experimental group and 89 percent for both the Section 8 and control groups.[21] All statistical estimates in this paper use weights.[22] The baseline covariates

---

[18] 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.

[19] The correlation in the amount reported for 2001 labor earnings on the survey and the total UI earnings in 2001 was .74 among individuals with valid reports for both (including zeros). The kappa statistic for agreement between reporting positive labor earnings in 2001 and UI earnings in 2001 was .63. While self-reports may have more recall error, they can also capture legitimate employment not covered by the UI system, such as domestic work, self-employment, federal work (civilian and military), and out-of-state employment. One anomaly in the survey data was a surprising lack of correlation between survey earnings collected by one particular interviewer and the administrative earnings data for the same individuals. In addition, the blood pressure data collected by this interviewer had a disproportionate share of values at exactly 120/80. Callbacks to these respondents indicated that the interviews did take place, but apparently had a great deal of measurement error. The 113 interviews conducted by this interviewer were spread across the three MTO groups in roughly equal numbers. All of the results reported in this paper include data collected by this interviewer, but we have replicated all analyses without this interviewer and the results are essentially unchanged.

[20] An initial phase from January - June 2002 resulted in an 80% response rate. 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 in our analyses. We calculate the effective response rate as $80 + (1 - .8)*48 = 89.6$. The full sample consisted of one adult from each of 4,248 households; the total number of completed surveys was 3,526.

[21] The potential for non-response bias and other internal validity issues are discussed below in section III.G.

(X) included in the regressions in this paper are described in Table 1.   Supplemental details

regarding health outcome variables are given in the Appendix.

II. Baseline Characteristics, Mobility, and Neighborhood Context

The adults in our sample are mainly female and black or Hispanic, as shown in Table 1.[23]

93 percent were ages 25-54 as of December 31, 2001. At the time of random assignment, 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 72 comparisons using the

covariates in Table 1, there were two statistically significant treatment-control differences (at the

5 percent level).[24]

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

---

[22] The weights have two components. First, the 3-in-10 subsample of hard to find cases were weighted by 10/3.
Second, observations are weighted by the inverse of the random assignment ratio. This ratio was changed during the
course of the demonstration to minimize minimum detectable effects after take-up of the vouchers turned out to be
much higher than had been projected. The second component of the weights is, therefore, necessary to prevent time
or cohort effects from confounding the results. 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 second component
of the weights. See Appendix B of Orr et al. (2003) for additional details.

[23] For 91 percent of households, there was a female head and no spouse, and in 2 percent there was a male head and
no spouse -- and the sample adult was the head. In the remainder of households, with a head and a spouse, the
sample adult is the female among the two.

[24] The fraction that had moved three times in the five years prior to baseline was lower in the experimental than
control group. The fraction who had previously applied for Section 8 was higher in the Section 8 group than the
control group. An omnibus F-test of the differences between the treatment groups and the control groups can be
constructed using the seemingly unrelated regression system in equation (6), with the 36 covariates in Table 1 (plus
fifth order polynomial in age) stacked as Y and with X containing only a treatment indicator and a constant. The p-
value on the joint significance for the experimental vs. control differences was .67, and for the Section 8 vs. control
differences was .64. These results are consistent with successful implementation of random assignment. We have
also compared the baseline characteristics of survey respondents with those of the full sample, with results shown in
the web appendix described in Appendix Table A0. The higher survey response rate for Chicago and the racial
composition at the Chicago site led to the survey sample in the experimental and Section 8 groups having a higher
fraction of Chicago residents and African-Americans than the full sample. Of the 102 respondent versus full sample
comparisons not involving site or race variables, three had statistically significant differences.

the Section 8 group.[25] Compared to non-compliers, compliers 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.[26] In multivariate estimation of equation (1), age and dissatisfaction with

apartment are significant predictors of compliance for both treatment groups; optimism about

finding a new apartment and concurrent school enrollment also are significant predictors of

compliance in the experimental group. 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 LA Section 8

group.[27]

To assess the impacts of the randomized offer of a voucher on residential location, we

present ITT, TOT, and CCM estimates from equations (2) - (4) of effects on selected residential

neighborhood and housing characteristics. In the first row of Table 2, the average census tract

poverty rate for the control group was .45. This poverty rate is calculated by averaging the

poverty rates from each address at which the individual lived between random assignment and

2002, weighting by the duration of residence at that address.[28] Annual poverty rates for a census

tract are obtained by linear interpolation from the rates estimated in the 1990 and 2000 censuses.

---

[25] These compliance rates are calculated using the sample weights and, therefore, differ slightly from the raw sample counts shown in Table 1.

[26] The decision to use an MTO voucher was also likely influenced by the limited period (typically 120 days) in which the voucher had to be used after it was offered, the availability of units on the private housing market, and the concurrent Hope VI renovation of many of the original housing projects which made them more attractive.

[27] These rates differed primarily because the intensity of housing search assistance provided to experimental group households by the non-profits responsible for the counseling varied considerably across sites (Feins, McInnis, and Popkin, 1997). The tightness of local housing markets presumably mattered as well.

[28] We view this average poverty rate as purely a descriptive measure. It is possible that some other feature of people's mobility history -- such as the minimum or maximum quality neighborhood to which they are exposed -- is more relevant for outcomes.

The intent-to-treat estimates indicate that average tract poverty was .12 lower for the experimental group and .10 lower for the Section 8 group.[29]

The differences between the tract poverty rates for treatment compliers and the implied distribution for control compliers are starker. Assuming no effect on those who did not use an MTO voucher to move, the treatment-on-treated effect in Table 2 for the experimental group was a reduction in average poverty rate of .26, and a reduction of .16 for the Section 8 group. The marginal distributions for these groups, based on kernel density estimation, are shown for the experimental group in Figure 1 and for the Section 8 group in Figure 2.[30] About 13 percent of the control complier density lies below 30 percent poverty in both the experimental and Section 8 groups, as can be seen by integrating over the densities in Figures 1 and 2 or calculated using equation (4) with the resulting control complier means shown in columns 4 and 8 of the second row of Table 2. In contrast, about 87 percent of the experimental compliers and about 53 percent of the Section 8 compliers had average tract poverty rates of less than 30 percent.[31]

Further mediating factors -- selected to be illustrative -- are shown in Table 2.[32] In addition to having lower poverty rates, the census tracts where treatment group members resided

---

[29] The differences are largest soon after random assignment and then decline over time. There are three main reasons for this. First, the census tracts to which MTO compliers initially moved had higher poverty rates on average in 2000 than in 1990. Second, subsequent moves by treatment compliers tended to be to tracts with higher poverty rates. Third, many control group families eventually left their original housing projects and these moves tended to be to census tracts with lower poverty.

[30] The overall distribution for the treatment group (f_1) can be decomposed into the fraction (π_1) due to the density for compliers (f_{11}) and the fraction ($1-\pi_1$) due to the density for noncompliers (f_{10}), where all three distributions are directly observable. For controls, only the overall distribution (f_0) is observed. Imbens and Rubin (1997) show that under the assumption that treatment noncompliers have the same distribution of poverty rates as the control noncompliers ($f_{10} = f_{00}$), one can subtract the treatment noncomplier density from the control overall density to obtain an estimate of the control complier density: $f_{01} = (f_0 - (1-\pi_1)f_{10})/\pi_1$. The kernel estimates used in Figures 1 and 2 are based on an Epanechnikov kernel with a halfwidth of .020.

[31] The treatment complier means are observed and can be calculated directly. They can also be recovered from TOT and CCM estimates in Table 2 using equation (4), and from integration over the densities in Figures 1 and 2. Note that equation (4) uses regression adjustment and the density estimation does not -- so control complier mean estimates from these two calculations are very similar but not identical.

[32] A more complete set of neighborhood contextual factors are summarized in Table 7 and discussed in section III. Results for additional specific items about mobility and neighborhoods are given in the web appendix as described in Appendix Table A0.

had a lower share of the population on public assistance and a greater share employed and in professional and managerial occupations than did the tracts of control group members. All of these effects were larger for the experimental group than for the Section 8 group. The fraction witnessing drug selling in their neighborhoods was .46 in the control group. The estimated effects of the treatment on the treated were quite large: a .25 reduction for the experimental group and a .17 reduction for the Section 8 group, relative to control compliers.

The treatment had relatively modest impacts on the average census tract percentage minority, and few treatment group compliers moved more than 10 miles from their origin location. Relative to the control group, the experimental group's housing units were somewhat higher quality as measured by problems with vermin, while the difference for the Section 8 group was negligible. Measures of social networks similarly show modest impacts overall. While experimental group compliers were about 20 percent more likely than controls to have a friend who is a college graduate or earns more than $30,000, the effect for the Section 8 group was not statistically different from zero, and other social network measures such as the frequency of church attendance showed no significant effects.

III. Results

The MTO demonstration caused a large fraction of those who used an MTO voucher to live in neighborhoods quite different from those in which they would have lived had they not been offered a voucher. A rough characterization of the results in Section II is that the housing vouchers were used to move to neighborhoods that were substantially less poor and less distressed, but not to move to high quality housing in racially integrated neighborhoods. This section examines the effect of using an MTO voucher on the domains of economic self-

sufficiency, mental health, and physical health. To focus our discussion, we first present results for five key measures in each domain, followed by results for summary measures, subgroups, and mediating factors.[33] The section ends with a general discussion of internal and external validity issues.

A. *Economic Self-sufficiency*

We hypothesized that moving to neighborhoods with lower poverty rates and higher rates of employment would provide MTO movers with stronger networks for job seeking, stronger community norms supportive of work, and greater physical proximity to employment opportunities. If there were improvements in mental and physical health, say from living in safer and less stressful environments, we anticipated that these could also contribute to greater employment and reduced public assistance use in the treatment groups. Countervailing factors were thought to be disruptions, particularly in the short run, in social networks that provide information about employment opportunities and reductions in the availability of informal child care. Previous evidence has generally suggested that living in a disadvantaged neighborhood reduces work, although early studies of specific MTO sites found no effects on employment within the first several years after random assignment.[34] An important motivation for longer follow-up was the research on central city versus suburban moves based on the Gautreaux program, which found effects on employment for a sample observed an average of five years

---

[33] The three domains and the key outcomes within each domain were pre-specified as part of the study design. Results for additional outcomes, pre-specified in the study design as lower priority than the ones presented in the main results, are available in the web appendix as described in Appendix Table A0.

[34] Weinberg, Reagan, and Yankow (2004) analyzed how work hours are related to neighborhoods in the NLSY using individual fixed effects. Ong (1998) compared hours of work for Section 8 recipients and public housing residents. Regarding initial results on economic outcomes for specific MTO sites, see Ludwig, Duncan, Pinkston (2005); Katz, Kling, and Liebman (2001); Hanratty, McLanahan, and Pettit (2003); and Leventhal and Brooks-Gunn (2003b).

14

after their initial move -- approximately the same average duration as the MTO survey data we examine.[35]

The results for employment, earnings, and public assistance receipt for adults in all five MTO sites, using the respondents' self-reports from survey data, are shown in Table 3. 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.[36] The experimental and Section 8 groups were slightly more likely to be employed, modestly less likely to be on TANF, and slightly more likely to be both employed and not on TANF than the control group. None of these treatment effects are statistically significant, although we do not have the statistical power to detect changes of less than 10 percent of the control group mean.[37] Furthermore neither the experimental nor the Section 8 group differed significantly from the control group in earnings in 2001 or in income from government sources in 2001. Additional results, not shown in the tables, for outcomes related more directly to well-being than to self-sufficiency per se, such as total income and food security, also show no statistically significant effects.

---

[35] Popkin, Rosenbaum, and Meaden (1993) analyzed suburban and central city residents in the Gautreaux program after five years using survey data, and Mendenhall, Duncan, and DeLuca (2003) followed-up longer-term outcomes with administrative data. Although there may be selection bias in these studies, the results when taken at face value suggest the possibility of substantial effects of residential location on employment. This was true even though the Gautreaux participants originally put low priority on jobs as a reason for moving (Peroff et al., 1979), just as MTO participants also put low priority on jobs as a reason for moving when asked at the time of random assignment.

[36] 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.

[37] The minimum detectable effect (MDE) that could be detected 80 percent of the time with 95 percent confidence is roughly 2.8 times the standard error of the estimate (Orr 1999). The MDEs are about 10-15 percent of the control mean for the ITT employment and earnings estimates in Table 3, and about 20-30 percent of the CCM for the TOT estimates.

Although we did not find differences between the groups on employment, the labor market context in which this study took place was one of dramatic change. The national unemployment rate fell from 6.1 percent in 1994 to 4.0 by 2000, before rising again to 5.8 in 2002.[38] There was a substantial increase in incentives for work from welfare reform and expansions of the Earned Income Tax Credit, and during the 1990s welfare rolls plummeted and employment rates for female heads rose substantially (Blank and Ellwood, 2002).[39] In addition, as time passed, the adults in our sample were less likely to have young children at home. As shown in Figure 3, employment rates increased for the MTO population over this period, rising from less than 25 percent in early 1995 to more than 50 percent in 2001.[40] However, the time-patterns appear similar for the three randomly-assigned groups.

This apparent lack of differences between the MTO groups in the administrative data is confirmed in panel A of Table 4 which shows administrative data employment results for three different time periods (fraction of quarters with positive UI earnings in 2001, in years 1 to 5 after random assignment, and in year 5 after random assignment).[41] Results for earnings in panel B

---

[38] The average of the unemployment rates for the five metropolitan areas in the MTO demonstration is slightly higher but tracks the same pattern (1994: 6.8; 2000: 4.3; 2002: 6.1).

[39] The decline in welfare income and increase in earnings had roughly offsetting effects on the consumption of single mothers in the late 1990s (Meyer and Sullivan, 2004).

[40] For our analyses based on UI data, "employment" is defined as positive earnings in a quarter, and individuals with no UI record in the quarter are assumed to have zero earnings. Note that by the end of 2001, approximately seven percent of our sample has at least one address outside of our five states with UI data. The administrative data understates the employment and earnings of those working in the informal sector or outside of their state of random assignment. For the four states with individual-level UI data, the fraction of quarters employed in 2001 is two percentage points higher when the sample is limited to those who never had an out-of-origin-state address, but the differences between MTO groups are virtually identical to those using the full sample.

[41] For simplicity and methodological consistency, the UI data from all of the states were aggregated into cells of at least 10 individuals and then analyzed at the cell level, to match the form of the data provided for Massachusetts. The same algorithm was used to construct the cells for all five states. The algorithm maintained the distinctions between sites, randomly assigned groups, and randomization periods. The cells were constructed to be as small as possible (but with at least 10) and to be as homogeneous as possible on randomization quarter, baseline education, and baseline work status. The cell-level analyses of mean outcomes control for these mean characteristics and for site. In constructing our administrative data outcomes, we collapse data from multiple quarters into a single observation per cell. With weights, the treatment and control groups are balanced on calendar time. The estimates that extend to five years after random assignment have smaller sample sizes than the calendar year 2001 estimates because sample members who were randomly assigned toward the end of the demonstration are not observed for

and for public assistance receipt and amount in panels C and D similarly fail to exhibit any statistically significant treatment effects.[42]

The neighborhoods into which treatment group families moved experienced changes during the 1990s that reduced differences between origin and destination neighborhoods and could possibly explain the lack of employment effects. The census tracts in which experimental group compliers lived in 2002 experienced a 13 percentage point increase in share minority, a 5 percentage point decline in share employed, and a 3 percentage point increase in the share poor between 1990 and 2000. Trends for the census tracts of Section 8 compliers were similar, and these trends for both treatment groups were significantly greater (in absolute value) than those for the control group.[43] Thus the families in the treatment groups moved into neighborhoods where the level of employment was substantially higher but where the trend from 1990 to 2000 in characteristics such as employment among residents was actually declining a bit faster than the trend for the control group.

If the change in employment among residents reflects the hiring patterns of accessible employers, then there may have been fewer job vacancies available to treatment group compliers

five full years in our administrative data. We have compared cell-level analyses of treatment effects for the four states with individual-level data; in all cases, the results are almost identical using the two approaches. A more complicated estimator combining estimates from the individual-level data (with a full set of covariates) for these four states and the cell data for MA gives very similar point estimates with standard errors that are about five percent smaller.

[42] We do not find effects on welfare in 2001 or five years after random assignment for analysis limited to Baltimore, indicating that the reductions in public assistance receipt for the experimental group relative to the control group in Baltimore during the first two years after random assignment found by Ludwig, Duncan, and Pinkston (2005) did not persist.

[43] Analysis of various census tract characteristics at one year after random assignment and at current census tract location at the time of the survey in 2002 are shown in the web appendices as described in Appendix Table A0, including 1980, 1990, and 2000 tract characteristics and their decadal differences. Nationally, the employment to population (e/pop) ratio for ages 16+ was .653 in 1990 and .639 in 2000, for a decline of .014 (Clark and Weismantle, 2003). The 2002 census tract locations for the control group had an average 1990 e/pop ratio of .433 and a 2000 ratio of .412, for a decline of .021. Among experimental complier tracts, the e/pop ratio declined from .583 to .533 between 1990 and 2000. The 2002 locations of the experimental group overall had a 1990 e/pop ratio of .499 and a 2000 ratio of .466, for a decline of .033. The difference-in-difference of -.033 - (-.021) = -.013 has a p-value of .002. There was also a larger decline in tract employment for the Section 8 group relative to controls, with a difference-in-difference of -.009.

in their new neighborhoods.[44] However, using establishment-based data on employment we find

that between 1994 and 2001, employment in control group zip codes increased by 5.0 percent,

compared with 4.9 percent in experimental group zip codes (p-value on contrast=.89), and 4.3

percent in Section 8 group zip codes (p-value on contrast=.51).[45] Thus, to the extent that

aggregate employment growth at the zip code level reflects available job vacancies, there do not

appear to be important differences across the three groups in job accessibility. We found no

significant between group differences on commuting-related outcomes.[46]

We also find some evidence that the social networks relevant to employment for this

population may not have been greatly affected by the change in residential location. In results

not shown in the tables, we find that only about .08 of the sample found a job through someone

living in their neighborhood such as a friend, relative, or acquaintance -- and that this proportion

did not vary across MTO groups.[47]

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 country – had no discernable

overall effects on employment, given that the most recent comprehensive survey (Ihlandfeldt and

Sjoquist, 1998) concludes that the empirical evidence overwhelmingly supports the spatial

mismatch hypothesis. Spatial mismatch has come to represent a collection of related ideas, and

the MTO experiment is informative about some of these but not others. There was no effect of

the experiment on neighborhood labor market conditions and job accessibility that would have

---

[44] Ihlanfeldt and Sjoquist (1998) argue that measures of job accessibility based on job growth do a better job of explaining spatial mismatch than measures based on the level of employment. See Raphael (1998).

[45] These estimates use the Census Zip Code Business Patterns data. Details are provided in the web appendices as described in Appendix Table A0.

[46] For example, 92 percent of adults in the experimental group and 90 percent of those in both the Section 8 and control groups said that they were within 15 minutes of the nearest train or bus stop. Forty percent of the experimental group, 41 percent of the Section 8 group, and 38 percent of the control group said that someone in their household "owns a car or truck that runs."

[47] The types of informal networks emphasized by Bayer, Ross, and Topa (2004) may exist for this sample, but do not seem to be of large importance.

been expected to have led to increases in employment, in that distance moved was not large and there was not an effect of moving to areas with greater job growth. That is, it turns out the experiment was not a test of this aspect of spatial mismatch. To the extent that the idea of spatial mismatch more broadly encompasses residence in distressed, unsafe, crime-ridden communities inhibiting access to jobs outside these communities, the MTO experiment does indicate that moving to communities that are substantially safer on average does not have detectable effects on employment.

In summary, we find little indication of significant overall MTO treatment impacts on adult economic self-sufficiency. There is some suggestive evidence of improvements in employment and earnings for the younger adults in the experimental group, discussed further below when we examine effects for subgroups.

B. *Mental Health*

Moving from a high-poverty neighborhood to a lower-poverty neighborhood is a major life event that may or may not improve health outcomes, including mental health. We speculated, on the one hand, that moving could result in social or cultural isolation and have adverse effects on mental health, and, on the other hand, that relocation to low-poverty neighborhoods could improve mental health outcomes by reducing exposure to violence.

The results for mental health outcomes are reported in panel A of Table 5. Overall, there is a consistent sign pattern of improvements in mental health on all five measures for both groups offered vouchers relative to the control group. The effects on psychological distress and on feeling "calm and peaceful" are statistically significant at the 5 percent level for the experimental group relative to the control group.

The results in column 3 show that the TOT effect was a reduction in distress of about two-tenths of a standard deviation.[48] While we decided in our ex-ante research design to report a scale score for distress in order make use of the full range of information in the scale, another way of assessing the magnitude of this effect is to examine the probability of being above a specific cutpoint on the scale. An unstandardized raw score of 13 or higher has been validated by in-depth psychiatric interviews as a useful cutpoint to screen for serious mental illness (SMI); roughly speaking, 20 percent of the U.S. population is estimated to have at least one psychiatric condition, and about one-third of these (7 percent) have symptoms severe enough to be classified as SMI.[49] In the MTO sample, the control complier mean for the fraction with SMI is .17, and there is a TOT effect of -.077 (p-value of .015), which is a relative risk reduction of 45 percent. In terms of other cutpoints, there are significant reductions in distress using raw scores of 6 or higher (CCM = .51) through 14 or higher (CCM=.11). Under an assumption of perfect rank correlation between treatment and control complier outcomes (Heckman, Smith, and Clements 1997), these results are consistent with significant reductions in distress within the more-distressed half of the complier population.

The most significant effect in Table 5 is on tranquility -- the increase in the fraction who report feeling calm and peaceful most or all of the time in the past 30 days (TOT=+.131; p-value=.006) in the experimental group relative to the control group. Table 5 also shows a

---

[48] The distress scale measures how much of the time during the past 30 days the sample member reports feeling: "so sad nothing could cheer you up," "nervous," "restless or fidgety," "hopeless," "everything was an effort," or "worthless." The scale was designed for the U.S. National Health Interview Survey to provide a national measure of psychological distress. For ease of interpretation, we have converted the standard K6 score, ranging from 0 to 24 with mean 5.78 and standard deviation 5.39, to a z-score.

[49] SMI is defined as having a non-substance-use psychiatric disorder in the past 12 months with a Global Assessment of Functioning score of less than 60. SMI prevalence is used in the allocation of block grant funds for community mental health services under the Alcohol, Drug Abuse and Mental Health Administration Reorganization Act. Based on administration of screening questions and followed up by detailed psychiatric assessment interviews, the probability that a randomly selected individual with SMI and a randomly selected individual without SMI will be correctly classified by a K6 raw score of 13 or higher is estimated to be about 85 percent. See Kessler et al. (2003) for details.

marginally significant TOT effect of a .059 reduction in probability of a diagnosis of a major depressive episode (p-value=.06; CCM=.196), which is a 30 percent reduction in relative risk. As perspective on the large magnitude of this reduction, we note that this effect is larger than many of the most effective clinical interventions for depression.[50] The effects on worrying and sleep are not statistically significant for the experimental group versus controls. No mental health effect estimates are statistically significant for the Section 8 group relative to the control group. Overall, these mental health findings for the full MTO sample of 3,526 adults observed on average 5 years after random assignment are broadly consistent with the early results for two of the MTO sites.[51]

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, they are unlikely to translate into effects on earnings that are large enough for us to detect.[52]

---

[50] For example, in a study often cited as an exemplar of an effective 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.

[51] In a sample of 511 in Boston two years after random assignment, Katz, Kling and Liebman (2001) found highly significant effects on feeling "calm and peaceful" for both the experimental and Section 8 groups -- effects larger than those reported here for the experimental group and found effects on depression for the experimental group with essentially the same relative risk reduction for compliers as those reported here, but which were not statistically significant. In a sample of 550 in New York three years after random assignment, Leventhal and Brooks-Gunn (2003a) used different scales of depression and anxiety without comparable magnitudes to those reported here and found statistically significant effects of better mental health among the experimental group (but not the Section 8 group) relative to the control group.

[52] One way of calculating the potential effect of mental health improvements on employment outcomes is to estimate a non-experimental regression of employment on mental health and other covariates and to use the coefficient on mental health as an indication of what would happen to employment if mental health were improved. Using equation (2) but substituting a mental health measure for Z and limiting the sample to the MTO control group, we found that depression was associated with 24 percentage points less employment, and that a unit rise in the distress z-score was associated with 12 percentage points less employment. The decrease in log odds of employment for the depressed relative to the non-depressed in the MTO control group was more than twice as large as estimates found in the literature for the general population, such as in Ettner, Frank, and Kessler (1997). Nevertheless, effects of mental health on employment as large as these would not be large enough to be reliably detected within the MTO research design. Multiplying the ITT effect by the non-experimental coefficient related to work, we find implied

C. *Physical Health*

We hypothesized that moving to new residential locations could decrease environmental irritants, provide safer locations for exercise, reduce stress, and increase access to more nutritious foods. Peers in lower poverty neighborhoods may have healthier eating and exercise habits. Health care access could also change. Effects on such mediating factors could therefore lead to effects on physical health outcomes.

Results for five physical health measures are shown in panel B of Table 5. The estimated effect is statistically significant for only one measure and the sign pattern of effects is not consistent across measures. Both voucher groups report worse overall health and have greater measured hypertension, but they report lower prevalence of asthma attacks, fewer problems carrying groceries or climbing stairs, and lower obesity. Of these results, only those for obesity are statistically significant.

Using a standard cutpoint of a body mass index (BMI) of 30 kg/m^2 or higher, we find that nearly half of the control group would be classified as obese.[53] The TOT estimate for the experimental group relative to the control group is a .103 reduction in the fraction obese (CCM=.502, p-value .03), a reduction in relative risk of about 20 percent. The TOT estimate for the Section 8 group is a .077 reduction (p-value .06).

effects on employment via reduced depression (.027*.24=.006) and distress (.092*.12=.01) that are one percentage point or less. Even if the experimental group voucher had caused all compliers to have no depression, the implied increase in the employment rate would have been less than 2 percentage points (.47*.196*.24=.022), which would be approximately the same size as the standard error on the ITT employment estimates.

[53] For example, BMI=30 corresponds to being five feet four inches tall and 175 pounds. Nationally for ages 18-44, 33 percent of African-American women and 22 percent of Hispanic women have BMI $\geq$ 30 (Lucas, Schiller, and Benson 2004). Our data on weight are self-reported. Studies obtaining both self-reported and directly measured data report small but systematic biases in over-reporting height and under-reporting weight that lead to a small downward bias in BMI (Engstrom et al. 2003). Females in the U.S. underestimate their weight by about 1.5 kg, and a desire to weigh less and a perception of being overweight are both related to the magnitude of underreporting (Villanueva 2001). Changes in neighborhood could affect reporting of weight in two offsetting ways. First, movers might perceive themselves as heavier relative to those around them and over-report their weight. Alternatively, movers might feel more self-conscious about their weight in new neighborhoods and under-report.

D. *Internal and External Validity*

Regarding the internal validity of these results, two key concerns include the use of self-reports and the possibility of attrition bias. All of the health measures other than blood pressure were self-reported, and neither the participants nor the interviewers were blinded to the intervention. Thus, it is possible that the estimated health impacts are due to some sort of reporting bias. However, the consistency between survey and administrative self-sufficiency estimates discussed above and the negligible estimates of treatment effects for many outcomes help rule out the most obvious types of reporting bias.[54] 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. 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.

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

---

[54] Other work on youth arrest records in MTO also finds no evidence of respondents in households offered vouchers reporting more socially desirable outcomes (Kling, Ludwig, and Katz, 2005). 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. It is also possible that the 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.

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.[55] 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.[56]

Regarding external validity, 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. Our analysis is informative about the families that participated in the MTO demonstration but does not capture potential spillover effects on residents of the neighborhoods to which voucher recipients move or from which they left. Our results are most applicable to situations where the number of households moving into any particular neighborhood is small (as it was in MTO), as larger numbers may have different "endogenous" or "general equilibrium" effects (Manski 1993, Heckman 2001).

---

[55] 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.

[56] 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.

noneE. *Summary measures of main outcomes*

In this subsection, we aggregate the information for specific outcomes into summary measures for the domains of economic self-sufficiency, mental health, and physical health in order to examine the general pattern of results. We interpret these summary measures for each domain as indices 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.[57] The mean effect sizes, shown in the first two columns of Table 6, are based on equally weighted standardized treatment effects as described in equation (5). There are no statistically significant mean effects on any domain for the Section 8 group, and no significant mean effects on economic self-sufficiency or physical health for the experimental group. There is a beneficial effect of being offered an experimental group voucher on mental health.

The increase in the mean effect size for mental health in the experimental group is .084 standard deviations, shown in column 1 of Table 6, and has a p-value of .0052.[58] As we mentioned above, a principal reason for reporting summary measures is to focus on a small number of pre-specified statistical estimates for which we can assess the joint probability under

---

[57] For the fifteen key outcomes in our analysis, 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 specified. 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.
[58] An alternative would be to calculate a Wald test statistic on the joint significance of the five effects within mental health, say testing H_0: $\pi_{21}=\pi_{22}=\pi_{23}=\pi_{24}=\pi_{25}=0$. The Wald test has a p-value of .07 for the mental health domain. Note that the Wald test has the property of being more likely to reject the null when positively correlated outcomes (such as obesity and hypertension) have treatment effects of opposite sign than the same sign for any given magnitude of the effect sizes, due to the shape of the 95 percent confidence region. For the mean effect size test, treatment effects of opposite sign cancel each other out, while the power of the test is allocated to detecting consistency of effects. Consequently, the nondirectional Wald test has lower power to detect consistently beneficial effects than the mean effect size test we prefer. Another possibility would have been to have used a multivariate test for one-sided alternatives (e.g. Follman 1996), but we made an ex-ante decision to retain the ability to detect effects that were consistently beneficial or consistently adverse.

the null hypothesis of no true effects of observing at least one effect as large as those observed. Under a null hypothesis of no effect for all of the eight measures for the full sample (in the first two columns of Table 6) and if the estimates were independent, we would expect to observe at least one effect size among eight this large in magnitude about four percent of the time: $1-(1-.0052)^8 = .041$.[59] Using a modified Westfall-Young method (described in detail by Kling and Liebman, 2004) to account for dependence among the estimates, we calculate that the probability is .029 of observing at least one of eight effect sizes as large as that observed for mental health under the null hypothesis of no effect. Thus, even under this stringent familywise error control, we find that the effects on mental health are unlikely to be due to chance.

One way to integrate the mental health results for the experimental and Section 8 groups is to use the fact that some sites (such as LA) had larger effects on average poverty rates for both treatment groups, and other sites (such as Boston) had smaller effects. In a simple model where the effect on outcomes is linearly related to the poverty rate, we would expect larger effects on outcomes for treatment sites with the largest effects on poverty rates. To examine this relationship, Figure 4 shows a scatterplot of the mean effect sizes for mental health (on the y-axis) and poverty rate effects (on the x-axis) for each site and treatment group.[60] The mean effect size estimates come from equation (6), with Z expanded to contain the treatment indicator

---

[59] The Bonferroni-adjusted p-value would be .0416 (8 x .0052). Bonferroni adjustments are too conservative (produce p-values that are too high) when estimates are correlated, hence our preference for the modified Westfall-Young approach which accounts for the dependence among the estimates.

[60] If there were only one outcome on the y-axis (instead of a mean effect size for five outcomes), the points in Figure 1 would be the same as plotting the reduced form coefficients of a first stage (instruments on poverty rate) against those of a second stage (instruments on outcomes) from a model using the site-by-treatment indicators to instrument for the poverty rate. This graphical approach is similar to that used by Holzer, Katz, and Krueger (1991). Related work (Liebman, Katz, and Kling 2004) has examined such models in more detail, which essentially fit a line through the points in Figure 4 after removing main effects for each site. We interpret the poverty rate as an index of neighborhood attributes, and not as the effect of a change in the poverty rate holding other neighborhood attributes constant.

26

interacted with five site indicators. Since the effects are all relative to the control group, the figure is normalized so that the control group is at (0,0) for each site.

Figure 4 shows that greater reductions in poverty rates are associated with greater improvements in mental health. This pattern of results suggests that it was the characteristics of the neighborhoods to which people moved, rather than the moves per se, that produced the mental health gains.[61] 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. Since there are virtually no outcomes or mediating factors for which the Section 8 assignment was more beneficial than the experimental assignment, it is unlikely that there was an offsetting benefit of Section 8 on some other unmeasured dimension.[62] Instead, our interpretation is that the information provided by the counseling combined with the restriction that forced individuals to search for housing outside areas previously known to them seems to have led to benefits for the individuals upon whom the restriction was imposed.

F. *Summary measures of outcomes for subgroups*

We have also explored whether the effects differ by baseline characteristics, such as education level and gender composition of children in the household. In general, results do not differ appreciably by these characteristics.[63] Similarly, we find that our main conclusion (mean

---

[61] In theory, the intervention could also have operated by changing the amount that could be spent on non-housing goods, due to a change in out-of-pocket housing expenses. In practice, there is little difference between the housing costs of the two groups. Out-of-pocket housing costs per month including utilities (and including mortgage and insurance for homeowners, and zero costs for the homeless or institutionalized) averaged $395 for the experimental group, $386 for the Section 8 group and $375 for the control group.

[62] The two treatment groups had different compliance patterns. This makes it difficult to compare the experimental and Section 8 treatment effects because the people who were affected may have been different. For example, the compliance rate was higher for the Section 8 group, and it is possible that individuals who would have complied under both treatment offers actually had equivalent effects or even better effects from Section 8, but that treatment effects were smaller for those in the Section 8 group who would have complied only with the Section 8 offer.

[63] These results are given in the web appendix described in Appendix Table A0. Kling and Liebman (2004) examine the parents of youth ages 15-20 in MTO and find that the beneficial experimental group mental health

effects only for the experimental group on mental health) holds when the sample is restricted to minority female heads of household ages 25-54. Results do, however, differ noticeably for younger versus older adults. Examining effects by age was an exploratory and not a confirmatory exercise, designed to supplement our results after analyzing the summary measures for the full sample in the first two columns of Table 6. A set of subgroups over which to search was not pre-specified; 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.

In columns 3 - 6 of Table 6 we report the mean effect size estimates for those ages 32 or younger at random assignment (where age 33 is the median) and for those ages 33 or older, with the differences between effects for these two subgroups reported in columns 7 and 8.[64] We find a positive and significant impact of the MTO experimental treatment on our summary measure of overall (economic and health) outcomes for the younger adults and no significant overall impact for older adults. In terms of individual outcome domains, we find the effects on mental health did not differ appreciably by age, although the effects for each subgroup are of course less precisely estimated than the effects for the full sample. The younger adults in both the experimental and Section 8 groups had significantly positive mean effects on physical health that were significantly different from the negative mean effect 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. We speculate that the habits and behaviors of younger adults were more malleable and therefore more susceptible to a change in residential environment.

---

effects are concentrated among the parents of females. However, we do not find that the mental health effects vary by gender composition of children for the full sample of adults examined here.

[64] The estimates for the 15 individual outcomes are shown in Appendix Table A1.

A similar pattern of relatively better outcomes for younger adults holds for economic self-sufficiency in the experimental group, although these results are not statistically significant. In results not shown in Table 6, we find some suggestive evidence using UI data of interesting dynamics in the treatment effects on employment and earnings for the 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.[65] This pattern of results for the younger adults fits with initial disruptive effects of moves on employment combined with a longer-term assimilation effect of taking advantage of new labor market opportunities. Measured by calendar year rather than years since random assignment, the treatment effects on employment and earnings improved from 1998 to 2000 for the younger experimental group and then declined as the recession hit in 2001. A similar pattern is not apparent for the older adults.

G. *Summary measures of mediating factors*

In Section II, we examined selected characteristics of the neighborhood context, showing large changes in poverty rates and neighborhood drug activity and more modest changes in the racial composition of neighborhood and the distance moved. In Table 7, we summarize a comprehensive set of variables describing mediating factors, using the same mean effect size metric as in Table 6.[66] Our intention in collecting extensive data on mediating factors was to explore mechanisms through which changes in residential location might impact our primary economic and health outcomes.

---

[65] The impacts of MTO on employment and earnings by age group using the administrative data are summarized in Appendix Tables A2 and A3 respectively. The administrative data results by age group use only the 4 sites for which we have individual-level administrative UI data. We have also examined the cell data for MA using interactions of fraction age<33 with treatment indicators. The point estimates suggest that there are also differences in treatment effects by age in the MA data, but these estimates are quite imprecise.
[66] Results for each component of the summary measures of contextual factors are given in the web appendix described in Appendix Table A0.

Relative to the standard deviation in the control group, the changes in tract characteristics (poverty rate, employment rate, share on welfare, and share of managerial and professional workers) are very large -- .6 standard deviations for the experimental group. There are also large effects for neighborhood conditions and safety and smaller effects on housing quality. In terms of other factors hypothesized to be related (although not necessarily causally) to self-sufficiency and health outcomes, we find no effects on social networks, education and training, smoking and drinking, or health care access. We do find positive effects on exercise and nutrition in the treatment groups relative to the control group, consistent with the effects on obesity for those groups.

Given the large differences in physical health outcomes by age found in Table 6, we investigated whether there were differences by age in mediating factors that might be related to health outcomes. We found no age-related differences in health care access, exercise, or nutrition. While there were differences in effects by age on smoking and drinking (with older adults in the experimental group smoking and drinking more than controls), these were not found in both treatment groups as the differences in physical health were. There are significant differences by age in the mean effect size estimates for tract characteristics and safety. However, the mean effect sizes in these tables are based on intent-to-treat estimates and the compliance rate (shown in the first row of Table 7) is much lower for the older adults. TOT estimates (assuming no effects for those who did not use a voucher to move) for tract characteristics and safety do not exhibit any statistically significant age differences. Thus, among those who used MTO vouchers to move, the differences in physical health by age that we observed do not appear to be strongly related to any of our measures of neighborhood or social context.

IV. Conclusion

A randomized intervention that used housing vouchers to encourage families in high-poverty public housing to move to private rental units in low-poverty neighborhoods was associated with better mental health among adults in an experimental group relative to a control group. There were no discernable effects on summary measures of economic self-sufficiency or physical health for the full experimental group sample. There were also no significant effects on summary measures of self-sufficiency, mental health, or physical health in a group given traditional Section 8 vouchers.

The main strength of the study was its randomized design, which eliminates the selection bias that has made it difficult to interpret non-experimental analyses of the association between residential environment and individual outcomes.[67] Moreover, the intervention resulted in a relatively homogenous group of people living in a wide range of neighborhoods. The pattern of results -- with the agreement of estimated effects based on self-reports and administrative records, with effects concentrated in the single domain of mental health, and with mental health effect sizes systematically related to changes in neighborhood poverty rates -- is consistent with beneficial impacts on mental health of moving to less distressed neighborhoods.

We also found a statistically significant reduction in obesity for both treatment groups. Although there is a reasonable chance that the obesity results represent a false significance (type I error), it is worth noting that the fact that the intervention simultaneously improved mental

---

[67] The associations of neighborhood residence with the health outcomes of distress, depression, and obesity in this study are broadly consistent with other findings based on observational data (Ellaway, Anderson, and Macintyre, 1997; Yen and Kaplan 1999b; Silver, Mulvey and Swanson 2002). In contrast, we found no significant differences between the three study groups in employment and earnings or general health, although associations between neighborhood residence and these outcomes have been documented in observational studies (Ong, 1998; Robert, 1998; Kadushin et al., 1998; Duncan, Jones, and Moon, 1999; Ross and Mirowsky 2001).

health and reduced obesity is consistent with theories hypothesizing a link between the two conditions.[68]

Because the MTO intervention changed so many aspects of people's neighborhood and housing environment simultaneously, the research design was not ideal for determining the mechanisms that may have produced the improvements in mental health that we observed. Alternative interventions that change one feature of the environment at a time -- improving policing to increase safety or improving the physical condition of people's housing units without changing their locations -- would be more suitable for evaluating particular mechanisms. Nonetheless, based 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.[69]

From a policy perspective, drawing firm conclusions about the preferred approach to providing housing assistance would require data on many groups of individuals beyond the adult participants studied here.[70] Nonetheless, these results for adults suggest that interventions that improve distressed neighborhoods or assist people in leaving them can have important public health benefits. The consistent relationship between reductions in poverty and improved mental health that we documented and the fact that the neighborhoods to which families moved using traditional Section 8 vouchers were less distressed suggests that there were likely to have been some health benefits for the Section 8 group relative to the control group, though they were too

---

[68] Carpenter, Hasin, Allison, and Faith, 2000; Faith, Matz, and Jorge, 2002; Friedman and Brownell, 1995; Bjorntorp, 2001.
[69] Popkin, Harris and Cunningham, 2002; Kling, Liebman and Katz, 2004.
[70] These groups include neighbors in both the origin and destination neighborhoods as well as children in the MTO families.

small for us to detect. Further intervention in the form of intensive housing-search counseling or locational restrictions appears to be necessary in order to induce public housing residents to choose neighborhoods in which they will experience large improvements in mental health.

The combination of impacts on mental health and no impacts on earnings or welfare usage indicates that mental health outcomes may be more sensitive than economic outcomes to housing policies for low-income families, at least in the medium run. These adult mental health benefits may have important spillover benefits, particularly to children, since children have been found to have more problems in school and more behavior problems when their mothers are experiencing mental health problems.[71] Although public health concerns were a central motivation for the major U.S. housing legislation of the 1930s and 1940s, HUD does not currently emphasize health effects in discussion of housing policies.[72] The findings in this study may return health concerns to the more prominent place in housing policy that they held 60 years ago, with a new emphasis on the importance of mental health.

---

[71] For reviews of linkage between maternal mental health and child development, see Beck (1999) and Goodman and Gotlib (1999). Another possible spillover of improving mental health could be reduced Medicaid costs, although expenditures on mental health services are not large. Buck et al. (2001) estimate that 10 percent of non-elderly Medicaid expenditures are for mental health services, with average expenditures of about $2000 per year per mental health service recipient.

[72] For examples of the prominence of health concerns in the discussion of housing policy, see Roosevelt (1938) and Truman (1948). Outside of the half a percent of its budget dedicated to lead paint and other health initiatives, the current U.S. Department of Housing and Urban Development (2004b) budget summary does not mention potential health effects from the billions to be spent on "Promoting Decent Affordable Housing," "Strengthening Communities," or other initiatives.

APPENDIX: HEALTH MEASURES

This appendix provides details about the ten main mental and physical health measures used in this study. Results for these measures are given in Table 5.

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; Ronald C. 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 (Ellen E. Walters et al. 2003).[73]

---

[73] 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 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.

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).

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.

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, $BMI \geq 30 \text{ kg/m}^2$ (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.

REFERENCES

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.

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.

Bayer, Patrick; Ross, Stephen L. and Topa, Giorgio, 2004. "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes." Manuscript, Federal Reserve Bank of New York.

Beck, Cheryl H., 1999. "Maternal Depression and Child Behavior Problems: A Meta-analysis." *Journal of Advanced Nursing*, 29(3), 623-629.

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

Bjorntorb, Per, 2001. "Do Stress Reactions Cause Abdominal Obesity and Comorbidities?" *Obesity Review*, 2, 73-86.

Blank, Rebecca and Ellwood, David T., 2002. "The Clinton Legacy for America's Poor." In Jeffrey A. Frankel and Peter R. Orszag, eds., *American Economic Policy in the 1990s*. Cambridge, MA: MIT Press.

Bloom, Howard, 1984. "Accounting for No-shows in Experimental Evaluation Designs." *Evaluation Review*, 8, 225-246.

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.

Buck, Jeffrey A.; Teich, Judith L.; Bae, Jay and Dilonardo, Joan, 2001. "Mental Health and Substance Abuse Services in Ten State Medicaid Programs." *Administration and Policy in Mental Health*. 28(3), 181-92.

Carpenter, Kenneth M., Hasin, Deborah S., Allison, David B., and Faith, Myles S., 2000. "Relationships Between Obesity and DSM-IV Major Depressive Disorder, Suicide Ideation, and Suicide Attempts: Results from a General Population Study." *American Journal of Public Health*, 90, 251-257.

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.

Clark, Sandra L. and Weismantle, Mai, 2003. *Employment Status: 2000.* Washington, DC: U.S. Census Bureau.

Committee on Ways and Means, 2004. *The 2004 Green Book: Background Material and Data on Programs within the Jurisdiction of the Committee on Ways and Means.* Washington, DC: U.S. House of Representatives.

Cutler, David M. and Glaeser, Edward L., 1997. "Are Ghettos Good or Bad?" *Quarterly Journal of Economics*, 112(3), 827-872.

Diez-Roux, Ana V.; Merkin, Sharon Stein; Arnett, Donna; Chambless, Lloyd; Massing, Mark; Nieto, F. Javier; Sorlie, Paul; Szklo, Moyses; Tyroler, Herman A. and Watson, Robert L., 2001. "Neighborhood of Residence and Incidence of Coronary Heart Disease." *New England Journal of Medicine,* 345, 99-106.

Duncan, Craig; Jones, Kelwyn and Moon, Graham, 1999. "Smoking and Deprivation: Are there neighbourhood effects?" *Social Science and Medicine,* 48, 497-505.

Ellaway, A.; Andersen, A. and Macintyre, S., 1997. "Does Area of Residence Affect Body Size and Shape?" *International Journal of Obesity,* 21, 304-308.

Engstrom, Janet L.; Paterson, Susan A.; Doherty, Anastasia; Trabulsi, Mary and Speer, Kara L., 2003. "Accuracy of Self-reported Height and Weight in Women: An Integrative Review of the Literature." *Journal of Midwifery & Women's Health,* 48(5), 338-345.

Ettner, Susan L., Richard G. Frank, and Ronald C. Kessler, 1997. "The Impact of Psychiatric Disorders on Labor Market Outcomes." *Industrial and Labor Relations Review*, 51, 64-81.

Faith, Myles, S., Matz, Patty E., and Jorge, Marie A., 2002. "Obesity-Depression Associations in the Population." *Journal of Psychosomatic Research*, 52, 935-942.

Feins, Judith D.; McInnis, Debra and Popkin, Susan J., 1997. *Counseling in the Moving to Opportunity Demonstration Program.* Washington, DC: U.S. Department of Housing and Urban Development.

Follman, Dean, 1996. "A Simple Multivariate Test for One-Sided Alternatives." *Journal of the American Statistical Association*, 91(434), 854-861.

Ford, D. E. and Kamerow, D. B., 1989. "Epidemiologic Study of Sleep Disturbances and Psychiatric disorders." *Journal of the American Medical Association*, 262(11), 1479-84.

Friedman, Michael A. and Brownell, Kelly D., 1995. "Psychological Correlates of Obesity: Moving to the Next Research Generation." *Psychological Bulletin*, 117, 3-20.

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.

Goering, John M., 2003. "Comments on Future Research and Housing Policy," 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, 383-407.

Goodman, Sherryl H. and Gotlib, Ian H., 1999. "Risk for Psychopathology in the Children of Depressed Mothers: A Developmental Model for Understanding Mechanisms of Transition." *Psychological Review,* 10(3), 458-490.

Granovetter, Mark, 1974. *Getting a Job: A Study of Contacts and Careers.* Cambridge, MA: Harvard University Press.

Hanratty, Maria; McLanahan, Sara and Pettit, Becky, 2003. "Los Angeles Site Findings," 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, 245-274.

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

Heckman, James J.; LaLonde, Robert J. and Smith, Jeffrey A., 1999. "The Economics and Econometrics of Active Labor Market Programs," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics, Volume 3A.* Amsterdam: North Holland, 1865-2097.

Heckman, James J.; Jeffrey A. Smith; and Clements, Nancy, 1997. "Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts." *Review of Economic Studies*, 64(4), 487-535.

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

Holzer, Harry J.; Katz, Lawrence F. and Krueger, Alan B., 1991. "Job Queues and Wages." *Quarterly Journal of Economics*, 106(3), 739-768.

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(4), 849-892.

Imbens, Guido W. and Rubin, Donald B., 1997. "Estimating Outcome Distributions for Compliers in Instrumental Variables Models." *Review of Economic Studies,* 64(221), 555-575.

Ioannides, Yannis M. and Loury Linda D., 2003. "Job Information Networks, Neighborhood Effects, and Inequality." Tufts University Working Paper.

Jacob, Brian A., 2004. "Public Housing, Housing Vouchers, and Student Achievement: Evidence from Public Housing Demolitions in Chicago." *American Economic Review*, 94(1), 233-258.

Jargowsky, Paul A., 1997. *Poverty and Place*. New York: Russell Sage.

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.

Kadushin C.; Reber, E.; Saxe, L. and Livert, D., 1998. "The Substance Use System: Social and Neighborhood Environments Associated with Substance Use and Misuse." *Substance Use and Misuse,* 33, 1681-1710.

Kasarda, John D., 1989. "Urban Industrial Transition and the Underclass." *Annals Of The American Academy Of Political And Social Science*, 501(1), 26-47.

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

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.

Kling, Jeffrey R.; Liebman, Jeffrey B. and Katz, Lawrence F., 2004. "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, forthcoming.

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*, forthcoming.

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 Experiment." Princeton IRS Working Paper 493.

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 Pinkston, Joshua C., 2005. "Housing Mobility Programs and Economic Self-sufficiency: Evidence from a Randomized Experiment." *Journal of Public Economics*, 89, 131-156.

Mallar, Charles D., 1979. "Alternative econometric procedures for program evaluations: illustrations from an evaluation of Job Corps." *Proceedings of the Business and Economics Statistics Section of the American Statistical Association*, 317-321.

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

Mendenhall, Ruby; Duncan, Greg J. and DeLuca, Stephanie, 2003. "Neighborhood Resources and Economic Mobility: Results from the Gautreaux Experiment." Manuscript, Northwestern University.

Meyer, Bruce D. and Sullivan, James X., 2004. "The Effects of Welfare and Tax Reform: The Material Wellbeing of Single Mothers in the 1980s and 1990s." *Journal of Public Economics*, 88, 1387-1420.

Millennial Housing Commission, 2002. *Meeting Our Nation's Housing Challenges.* Washington, DC: U.S. Government Printing Office.

Montgomery, James D., 1991. "Social Networks and Labor Market Outcomes: Toward an Economic Analysis." *American Economic Review*, 81(5), 1408-1418.

Montgomery, James D., 1992. "Job Search and Network Composition: Implications of the Strength-of-Weak-Ties Hypothesis." *American Sociological Review*, 57(5), 586-596.

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.

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.

Ong, Paul, 1998. "Subsidized Housing and Work among Welfare Recipients." *Housing Policy Debate*, 9(4), 775-794.

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

Orr, Larry, 1999. *Social Experiments.* London: Sage Publications.

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.

Page, Marianne E. and Solon, Gary, 2003. "Correlations Between Sisters and Neighbouring Girls in Their Subsequent Income as Adults." *Journal of Applied Econometrics*, 10: 545-562.

Peroff, Kathleen A; Davis, Coteal L.; Jones, Ronald; Curtin, Richard T.; and Morans, Robert W., 1979. *Gautreaux Housing Demonstration: An Evaluation of Its Impact on Participating Households*. Washington, DC: 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.

Popkin, Susan J.; Rosenbaum, James E. and Meaden, Patricia M., 1993. "Labor Market Experiences of Low-income Black Women in Middle Class Suburbs: Evidence from a Survey of Gautreaux Program Participants." *Journal of Policy Analysis and Management*, 12(3), 556-573.

Quigley, John M., 2000. "A Decent Home: Housing Policy in Perspective." *Brookings Papers on Urban Affairs*, 1, 53-99.

Raphael, Steven., 1998. "The Spatial Mismatch Hypothesis and Black Youth Joblessness: Evidence from the San Francisco Bay Area." *Journal of Urban Economics*, 43, 79-111.

Robert, Stephanie A., 1998. "Community-level Socioeconomic Status Effects on Adult Health." *Journal of Health and Social Behavior,* 39, 18-37.

Roosevelt, Franklin D., 1938. "Letters to Several Governors on the Adoption of Housing Legislation, March 1, 1938." *The American Presidency Project*. Santa Barbara, CA: University of California Santa Barbara. (Accessed 10/26/04; http://www.presidency.ucsb.edu/ws/index.php?pid=15603).

Ross, Catherine E. and Mirowsky, John., 2001. "Neighborhood Disadvantage, Disorder, and Health." *Journal of Health and Social Behavior,* 42, 258-276.

Sanbonmatu, 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.

Silver, Eric; Mulvey, Edward P. and Swanson, Jeffrey W., 2002. "Neighborhood Structural Characteristics and Mental Disorder: Faris and Dunham revisited." *Social Science & Medicine,* 55, 1457-1470.

Topa, Giorgio, 2001. "Social Interactions, Local Spillovers, and Unemployment." *Review of Economic Studies*, 68, 261-295.

Truman, Harry S., 1948. "Special Message to the Congress on Housing, February 23, 1948." Independence, MO: Truman Presidential Museum and Library. (Accessed 10/26/04; http://trumanlibrary.org/publicpapers/index.php?pid=1401).

U.S. Department of Housing and Urban Development, 2000. "Economic Cost Analysis of Different Forms of Assisted Housing." Office of Policy Research and Development Issue Brief II, December.

U.S. Department of Housing and Urban Development, 2004a. *Congressional Justifications for 2005 Estimates*. Washington, DC: U.S. Department of Housing and Urban Development.

U.S. Department of Housing and Urban Development, 2004b. *Fiscal Year 2005 Budget Summary*. Washington, DC: U.S. Department of Housing and Urban Development.

Villanueva, Elmer V., 2001 "The Validity of Self-reported Weight in U.S. Adults: A Population Based Cross-sectional Study." *BMC Public Health,* 1:11. (Accessed 10/26/04; http://www.biomedcentral.com/1471-2458/1/11)

Waitzman, Norman J. and Smith, Ken R., 1998. "Phantom of the Area: Poverty-area Residence and Mortality in the United States." *American Journal of Public Health*, 88, 973-976.

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.

Weinberg, Bruce A.; Reagan, Patricia B. and Yankow, Jeffrey J., 2004. "Do Neighborhoods Affect Hours Worked: Evidence from Longitudinal Data." *Journal of Labor Economics*, forthcoming.

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.

Wilner, Daniel M.; Walkley, Rosabelle P.; Pinkerton, Thomas C. and Tayback, Matthew, 1962. *The Housing Environment and Family Life: A Longitudinal Study of the Effects for Housing on Morbidity and Mental Health.* Baltimore: Johns Hopkins Press.

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

Wilson, William Julius, 1996. *When Work Disappears: The World of the New Urban Poor.* New York: Alfred A. Knopf.

Yen, Irene H. and Kaplan, George A., 1999a. "Neighborhood Social Environment and Risk of Death: Multilevel Evidence from the Alameda County Study." *American Journal of Epidemiology*, 149, 898-907.

Yen, Irene H. and Kaplan, George A., 1999b. "Poverty Area Residence and Changes in Depression and Perceived Health Status: Evidence from the Alameda County Study." *International Journal of Epidemiology,* 28, 90-94.

FIGURE 1. EXPERIMENTAL GROUP POVERTY RATE DENSITY

FIGURE 1. EXPERIMENTAL GROUP POVERTY RATE DENSITY
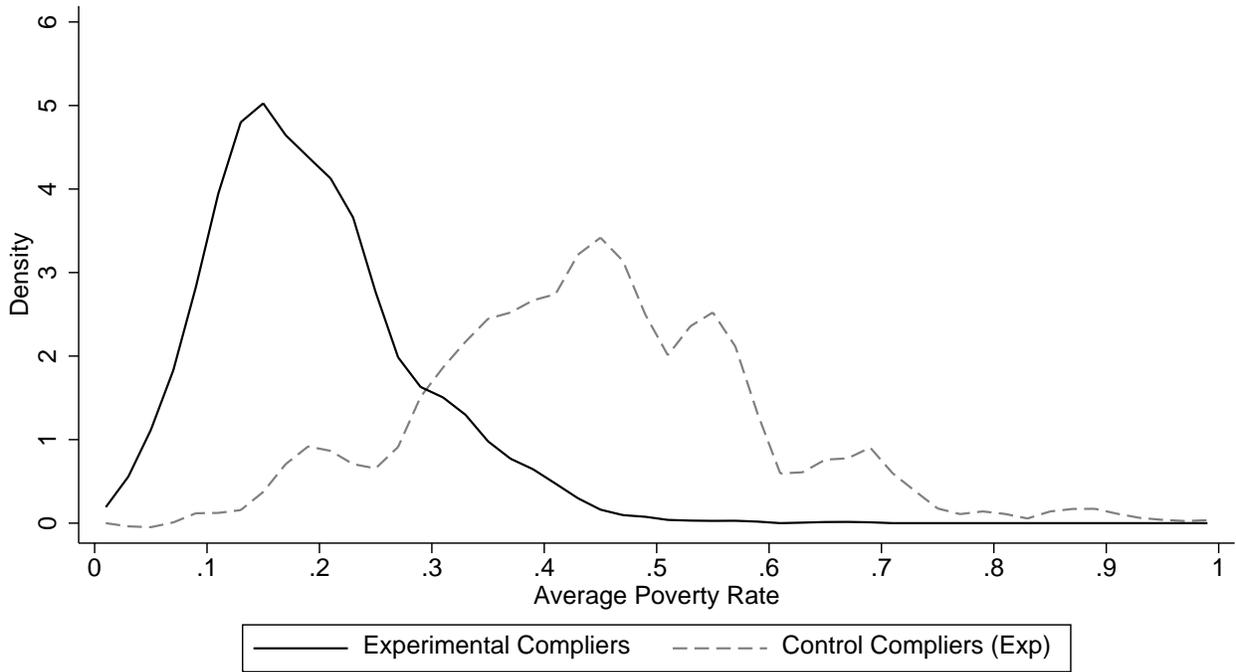
Notes. Kernel density based on actual distribution for experimental compliers and implied density for control compliers, as described in the text. Poverty rate averaged over tracts since random assignment, weighted by duration.

FIGURE 2. SECTION 8 GROUP POVERTY RATE DENSITY
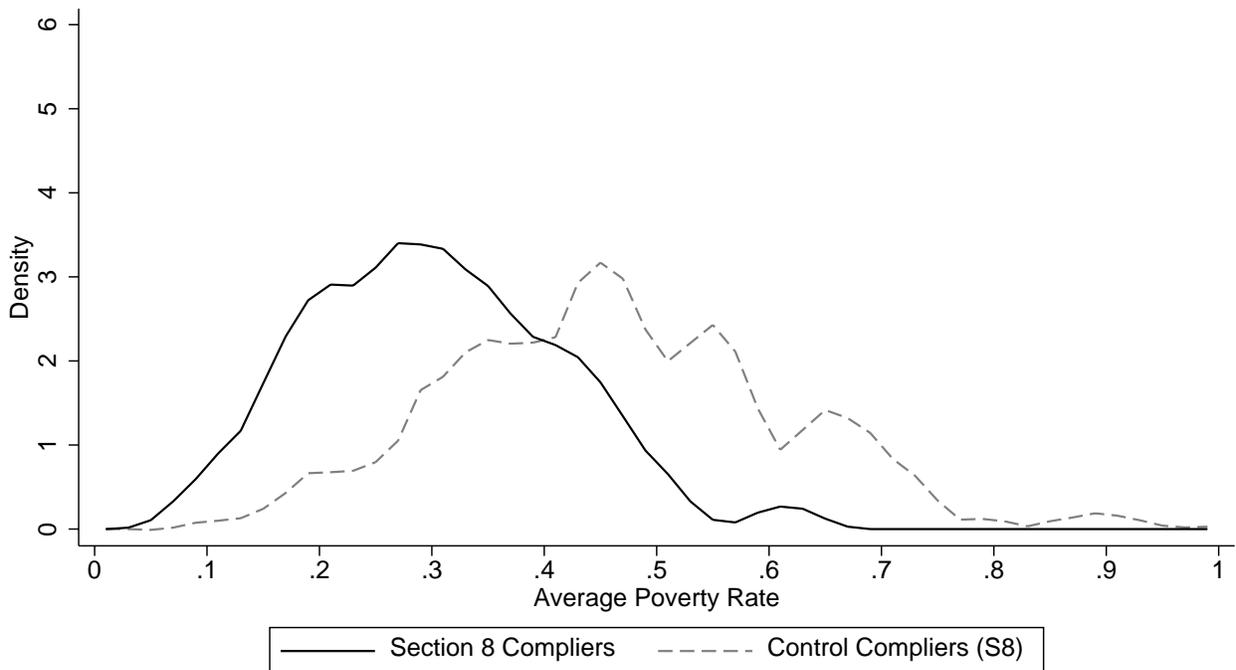
Notes. Kernel density based on actual distribution for Section 8 compliers and implied density for control compliers, as described in the text. Poverty rate averaged over tracts since random assignment, weighted by duration.

FIGURE 3. EMPLOYMENT RATES OVER TIME

FIGURE 3. EMPLOYMENT RATES OVER TIME
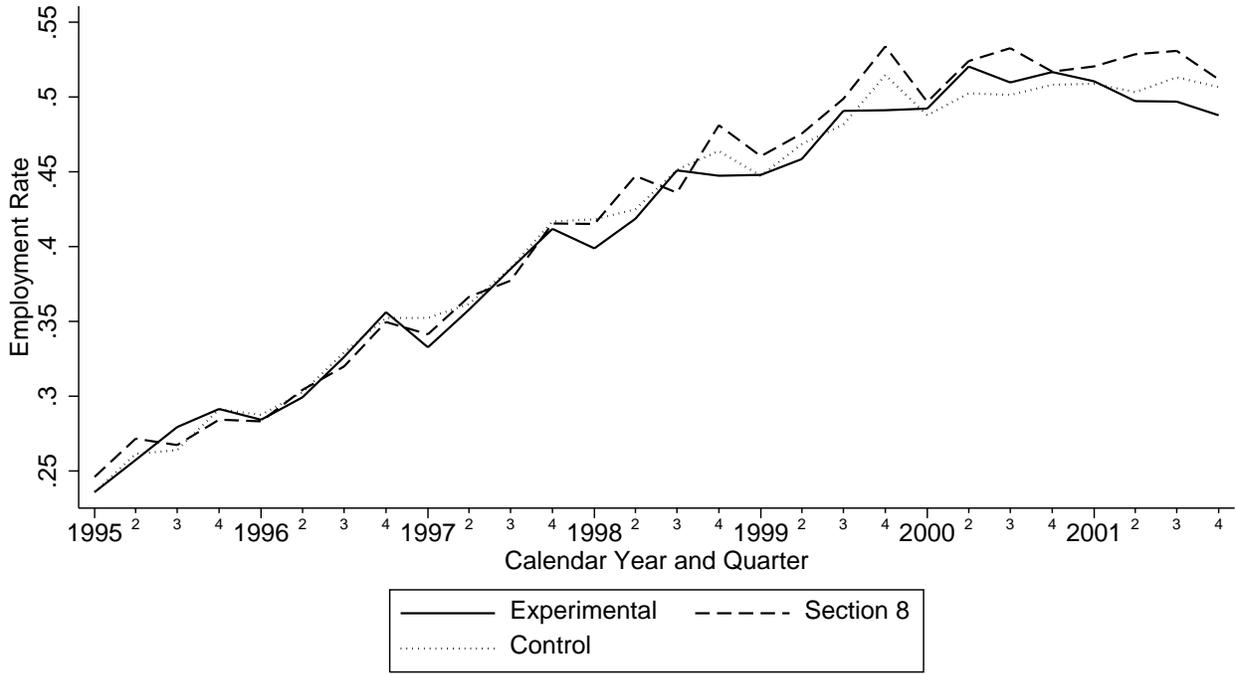
Notes. Employment is fraction with positive earnings per quarter from Unemployment Insurance records in California, Illinois, Maryland, Massachusetts and New York.

FIGURE 4. MEAN EFFECT SIZES FOR MENTAL HEALTH BY SITE
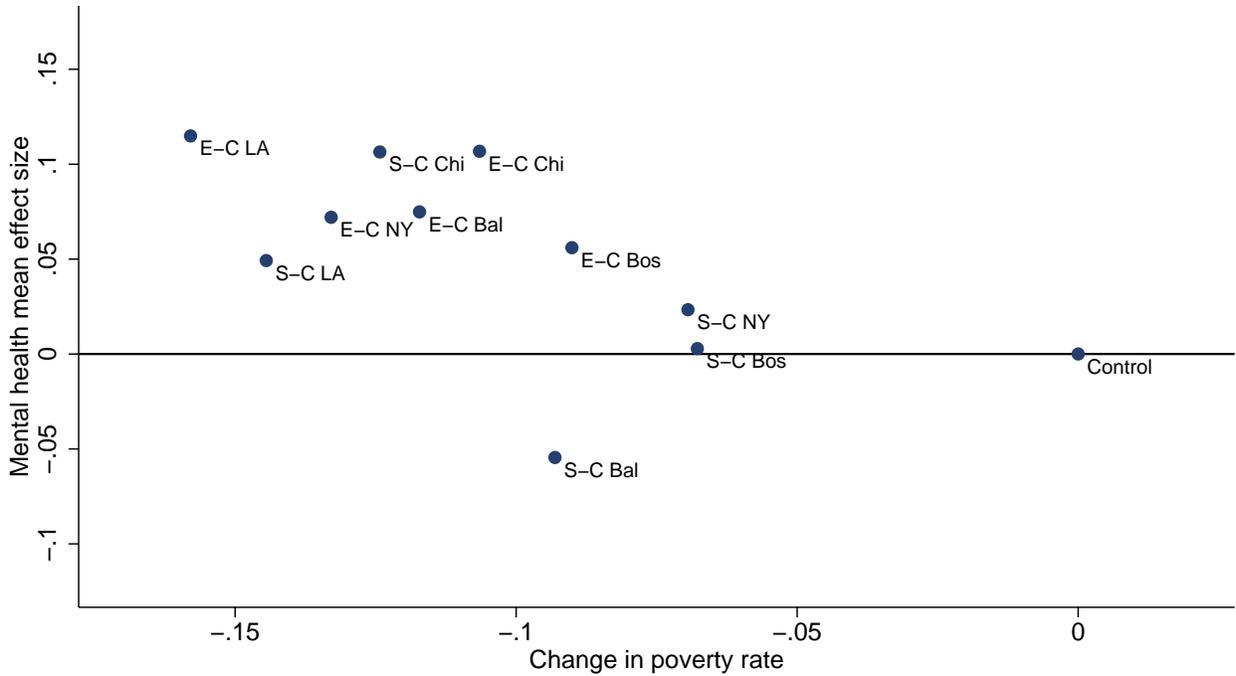
Notes. E-C: Experimental - Control. S-C: Section 8 - Control. Bal: Baltimore. Bos: Boston. Chi: Chicago. LA: Los Angeles. NY: New York. Mental health effect sizes based on equations (5) and (6), as described in the text. Poverty rate averaged over tracts since random assignment, weighted by duration, with estimates based on equation (2).

TABLE 1 ─ 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 1 ─ 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)
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. 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. * = 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 2 — EFFECTS ON SELECTED MEDIATING FACTORS

	CM	Experimental versus Control				Section 8 versus Control			
		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 (2), using covariates in Table 1 and weights described in section I. TOT = Treatment-on-treated from equation (3) 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 from equation (4). * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses.

TABLE 3 ─ EFFECTS ON ECONOMIC SELF-SUFFICIENCY

	CM	Experimental versus Control				Section 8 versus Control			
		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 (2), using covariates in Table 1 and weights described in section I. TOT = Treatment-on-treated from equation (3) 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 from equation (4). * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses.

TABLE 4 — EFFECTS ON EARNINGS AND WELFARE RECEIPT — ADMINISTRATIVE DATA

	CM	Experimental versus Control				Section 8 versus Control			
		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 (2), using covariates in Table 1 and weights described in section I. TOT = Treatment-on-treated from equation (3) 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 from equation (4). 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 5 ─ 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 (2), using covariates in Table 1 and weights described in section I. TOT = Treatment-on-treated from equation (3) estimated by two stage least squares with treatment group assignment indicator variables as the instruments for the treatment take-up indicator variables. * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses.

TABLE 6—MEAN EFFECT SIZES FOR SUMMARY MEASURES OF OUTCOMES

	All Adults		Age < 33 at RA		Age ≥ 33 at RA		Diff. by Age	
	E-C	S-C	E-C	S-C	E-C	S-C	E-C	S-C
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Self-sufficiency	.016	.034	.063	.047	-.032	.020	.095	.027
[5 measures]	(.031)	(.034)	(.046)	(.049)	(.043)	(.048)	(.063)	(.068)
Mental health	.084*	.030	.076	.028	.093*	.031	-.017	-.003
[5 measures]	(.030)	(.034)	(.043)	(.047)	(.042)	(.048)	(.060)	(.067)
Physical health	.016	.017	.072*	.080*	-.035	-.041	.107*	.121*
[5 measures]	(.024)	(.027)	(.035)	(.038)	(.035)	(.039)	(.049)	(.054)
Overall	.039	.027	.070*	.052	.008	.003	.062	.049
[15 measures]	(.020)	(.022)	(.029)	(.030)	(.029)	(.033)	(.041)	(.045)

Notes. RA = random assignment. E-C: Experimental - Control. S-C: Section 8 - Control. Estimates are the mean of the standardized intent-to-treat effects, from equation (5). Standard errors are derived from equation (6), adjusted for correlation within individuals. Economic self-sufficiency summarized the five outcomes in Table 3, reversing the sign of TANF and government income. Mental health summarizes the five measures in panel A of Table 5, reversing the signs of distress, depression, and worrying. Physical health summarizes the five measures in panel B of Table 5, reversing all signs. Overall summarizes all fifteen of these measures. The median age of the full sample at RA was 32.9 years old. Columns (iii) through (vi) summarize the estimated effects interacted with indicator variables for less than 33 years-old at RA and 33 or older at RA.

TABLE 7—MEAN EFFECT SIZES FOR SUMMARY MEASURES OF MEDIATING FACTORS

	All Adults		Age < 33 at RA		Age ≥ 33 at RA		Diff. by Age	
	E-C	S-C	E-C	S-C	E-C	S-C	E-C	S-C
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Used MTO voucher	.470	.598	.550	.697	.387	.501	.163*	.196*
Tract characteristics [4 measures]	.664* (.037)	.488* (.034)	.756* (.051)	.544* (.048)	.567* (.054)	.433* (.050)	.189* (.074)	.111 (.069)
Neighborhood condition [3 meas.]	.253* (.034)	.192* (.036)	.276* (.047)	.201* (.051)	.231* (.049)	.182* (.052)	.044 (.067)	.019 (.073)
Safety [5 measures]	.220* (.029)	.187* (.030)	.277* (.040)	.253* (.043)	.161* (.042)	.118* (.045)	.116* (.058)	.135* (.062)
Housing quality [4 measures]	.139* (.030)	.101* (.033)	.171* (.044)	.120* (.047)	.105* (.042)	.081 (.046)	.066 (.061)	.039 (.065)
Social networks [11 measures]	.007 (.016)	.022 (.018)	.012 (.023)	.027 (.026)	.003 (.022)	.019 (.025)	.009 (.032)	.008 (.036)
Education and training [3 meas.]	-.018 (.025)	.015 (.029)	-.003 (.036)	.039 (.041)	-.028 (.034)	-.009 (.041)	.025 (.050)	.049 (.058)
Exercise and nutrition [2 meas.]	.077* (.034)	.088* (.037)	.081 (.047)	.146* (.050)	.073 (.049)	.030 (.054)	.008 (.067)	.117 (.074)
Not smoking or drinking [2 meas.]	-.017 (.034)	-.017 (.037)	.048 (.049)	-.028 (.054)	-.099* (.049)	-.008 (.052)	.147* (.070)	-.020 (.075)
Health care access [2 measures]	.009 (.038)	.034 (.038)	-.021 (.059)	.013 (.060)	.036 (.047)	.050 (.048)	-.058 (.074)	-.037 (.076)

Notes. RA = random assignment. E-C: Experimental - Control. S-C: Section 8 - Control. Estimates are the mean of the standardized treatment effects from equation (5). Standard errors are derived from equation (6), adjusted for correlation within individuals. Signs for measures of adverse conditions are reversed such that positive mean effects indicate beneficial conditions; these measures are indicated below by: [-]. Tract characteristics, weighted by duration of residence since random assignment: poverty rate [-], fraction on welfare [-], fraction employed, fraction workers in managerial or professional occupations. Neighborhood quality: somewhat or very satisfied with neighborhood, fraction of reported problems [-] (trash on street, graffiti, public drinking, abandoned buildings, people hanging out), fraction of observed problems [-] (block poor condition, broken windows, metal bars, condition of street is poor, accumulation of trash). Safety: somewhat or very safe during day, somewhat or very safe at night, householder victimized [-], saw drugs sold [-], police not coming [-]. Housing quality: condition fair or poor [-], fraction of reported problems [-] (peeling paint, plumbing, vermin, roaches, broken locks, broken windows, heat), fraction observed interior problems [-] (noisy, cluttered, cracks, peeling paint, mold, smoke), fraction observed exterior problems [-] (unit exterior poor condition, broken windows, metal bars). Social networks: 3+ close friends, no friends in neighborhood [-], chat with neighbors, friend college graduate or earns $30K+, attends church 1+/month, trusts most people, experienced local discrimination [-], visited friends or relatives their homes, visited friends or relatives own home, found job through neighborhood contact, broad social network. Education and training: years of schooling, high school graduate, job training. Exercise and nutrition: days/week moderate exercise, days/week fruits or vegetables. Not smoking or drinking: smoker [-], binge drinking [-]. Health care access: has insurance, regular place to go for care.

APPENDIX TABLE A0. LIST OF APPENDIX TABLES

APPENDIX TABLE A1 — 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 (.030)	.021 (.033)	.439	-.013 (.028)	.010 (.031)	-.063 (.040)	-.011 (.045)
Employed [SR]	.555	.032 (.030)	.032 (.033)	.484	-.001 (.028)	.015 (.031)	-.033 (.041)	-.017 (.045)
Earnings in 2001 [SR]	9643	589 (659)	-508 (691)	7980	-362 (609)	486 (689)	-951 (896)	994 (980)
Receiving TANF [SR]	.334	-.036 (.027)	-.042 (.030)	.254	-.006 (.026)	-.020 (.028)	.030 (.037)	.022 (.040)
Income received from government sources during 2001 [SR]	2420	-84 (252)	-382 (277)	2552	479 (269)	179 (295)	563 (370)	561 (400)
B. Mental health								
Psychological distress, K6 z-score [SR]	-.021	-.090 (.064)	-.051 (.069)	.125	-.095 (.067)	-.012 (.075)	-.005 (.092)	.039 (.102)
Probability of major depressive episode [SR]	.153	-.021 (.020)	-.013 (.021)	.177	-.035 (.021)	-.014 (.024)	-.014 (.029)	-.001 (.032)
Worried, tense, or anxious [SR]	.360	-.015 (.030)	.026 (.033)	.429	-.043 (.031)	-.043 (.034)	-.028 (.043)	-.069 (.047)
Calm and peaceful [SR]	.474	.051 (.031)	.025 (.033)	.457	.073* (.031)	.003 (.035)	.022 (.044)	-.023 (.048)
Sleeps at least 7 and <9 hours per night [SR]	.479	.045 (.031)	.027 (.034)	.420	.020 (.032)	.005 (.035)	-.026 (.045)	-.021 (.049)
C. Physical health								
Has fair or poor health [SR]	.248	-.012 (.026)	-.030 (.028)	.416	.046 (.029)	.054 (.033)	.057 (.039)	.084* (.043)
Has trouble carrying groceries or climbing stairs [SR]	.332	-.038 (.029)	-.043 (.031)	.545	.001 (.030)	.002 (.033)	.039 (.042)	.045 (.045)
Had an asthma or wheezing attack during past year [SR]	.205	-.028 (.025)	-.031 (.026)	.221	.003 (.025)	.011 (.028)	.031 (.035)	.042 (.038)
Obese, BMI≥30 [SR]	.452	-.056 (.031)	-.069* (.034)	.484	-.040 (.032)	-.023 (.035)	.015 (.044)	.047 (.049)
Has hypertension, [M] SBP≥140 or DBP≥90	.227	-.030 (.027)	-.010 (.030)	.369	.075* (.031)	.055 (.035)	.104* (.041)	.064 (.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 1 and weights described in section I, 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.

	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 (2), using covariates in Table 1 and weights described in section I, 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.

	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 (2), using covariates in Table 1 and weights described in section I, 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.