

CENTER FOR HEALTH AND WELLBEING
PRINCETON UNIVERSITY
APRIL 2004

**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 *

* Princeton University and NBER, Harvard University and NBER, Harvard University and NBER, NBER.

Support for this research was provided by grants from the National Science Foundation to the National Bureau of Economic Research (9876337 and 0091854) and the National Consortium on Violence Research (9513040), as well as by the U.S. Department of Housing and Urban Development, the National Institute of Child Health and Development and the National Institute of Mental Health (R01-HD40404 and R01-HD40444), the Robert Wood Johnson Foundation, the Russell Sage Foundation, the Smith Richardson Foundation, the MacArthur Foundation, the W.T. Grant Foundation, the Spencer Foundation. Additional support was provided by grants to Princeton University from the Robert Wood Johnson Foundation and from NICHD (5P30-HD32030 for the Office of Population Research), by the Princeton Industrial Relations Section, the Bendheim-Thomas Center for Research on Child Wellbeing, the Princeton Center for Health and Wellbeing, and the National Bureau of Economic Research.

We are grateful to Todd Richardson and Mark Shroder of HUD; to Eric Beecroft, Judie Feins, Stephen Kennedy, Larry Orr, and Rhiannon Patterson of Abt Associates; to our collaborators Jeanne Brooks-Gunn, Alessandra Del Conte Dickovick, Greg Duncan, Tama Leventhal, Jens Ludwig, Bruce Psaty, and Robert Whitaker; to our research assistants Erin Metcalf and Josh Meltzer; to seminar participants at Princeton, Berkeley, and Michigan; and to numerous colleagues for valuable suggestions. Any findings or conclusions expressed are those of the authors.

© 2004 by Jeffrey R. Kling, Jeffrey B. Liebman, Lawrence F. Katz, and Lisa Sanbonmatsu. All rights reserved.

April 2004

**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

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. However, we find no significant overall effects on adult employment, earnings, or public assistance receipt -- though our sample sizes are not sufficiently large to rule out moderate effects in either direction. 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 treatment 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, but no significant effects on four other aspects of physical health (general health, asthma, physical limitations, and hypertension), and our summary measure of physical health was not significantly affected by the MTO treatment for the overall sample.

JEL classifications: I18, I38, J38.

Jeffrey R. Kling
Department of Economics
and Woodrow Wilson School
Princeton University
Princeton, NJ 08544
and NBER
kling@princeton.edu

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

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

Lisa Sanbonmatsu
1050 Massachusetts Avenue
National Bureau of Economic Research
Cambridge, MA 02138
lsanbonm@nber.org

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 argument for the potential linkage of an individual's economic prospects and their residence in a distressed community has been powerfully advanced by William J. Wilson (1987, 1996), and the importance of proximity to employment has its roots in the spatial mismatch hypothesis of John F. Kain (1968).¹ The location of one's residence may also affect the process of finding employment through the structure of local social interactions (Mark Granovetter 1974; James D. Montgomery 1991, 1992; Giorgio Topa 2001) and through the impacts of economic or racial segregation (Roland Benabou 1993, David M. Cutler and Edward L. Glaeser 1997).² A wide range of links between residential location and health outcomes have also been hypothesized (Ichiro Kawachi and Lisa F. Berkman 2000). 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 (Daniel M. 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. Meanwhile, non-experimental research has found strong associations between living in disadvantaged

¹ The spatial mismatch hypothesis states that inner-city low-skilled minority workers have relatively weak access to jobs because job opportunities are disproportionately in suburban areas and housing market discrimination plus commuting costs create barriers that prevent minorities from reaching the suburban jobs. See Keith R. Ihlanfeldt and David J. Sjoquist (1998) for a recent review of the spatial mismatch literature.

² William A. Brock and Steven N. Durlauf (2001) provide a broad overview of interactions-based models.

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 Public housing projects provide subsidized units for 2.5 million households, and Section 8 vouchers help another 1.5 million households rent private units.⁵ Housing assistance is rationed, so that only a minority of eligible families receive 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 distortions from providing public housing units instead of vouchers.⁶ To receive project-based

³ Examples include John D. Kasarda 1989; Paul A. Jargowsky 1997; Stephanie A. Robert 1998; Irene H. Yen and George A. Kaplan 1999a; Norman J. Waitzman and Ken R. Smith 1998; Ana V. Diez-Roux et al., 2001; A. Ellaway, A. Anderson, and S. Macintyre 1997; Irene H. Yen and George A. Kaplan 1999b; Catherine Ross and John Mirowsky 2001; and Eric Silver, Edward P. Mulvey and Jeffrey W. 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 Committee on Ways and Means (2000).

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; Edgar O. 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

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

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

income and the city's Fair Market Rent, set at the 40th percentile of area rents. See John M. Quigley (2000) and Edgar O. 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 (John M. Goering et al. 1999).

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. Counseling agencies also recruited landlords and facilitated leasing with landlords who might otherwise have been unfamiliar with the housing vouchers (Judith D. Feins, Debra McInnis and Susan J. Popkin, 1997).

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 the 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, examine a much larger sample, and are able to study much longer exposure to new residential environments. We also have collected more comprehensive measures of economic self-sufficiency, mental health, and physical health outcomes as well as a broader range of mediating factors to potentially illuminate the mechanisms by 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, asthma, physical limitations, and hypertension). The probability is reasonably high of observing one significant physical health result 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

⁹ Related MTO research on youth outcomes is available in Jeffrey R. Kling and Jeffrey B. Liebman (2004) and Jeffrey R. Kling, Jens Ludwig, and Lawrence F. Katz (2004).

¹⁰ Among analyses of adult outcomes, Jens Ludwig, Greg J. Duncan, and Joshua C. Pinkston (2004) examined Baltimore; Katz, Kling, and Liebman (2001) examined Boston; Maria Hanratty, Sara McLanahan, and Becky Petit (2003) examined Los Angeles; and Tama Leventhal and Jeanne Brooks-Gunn (2003a, 2003b) examined New York.

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

Let Z be an indicator for assignment to a voucher 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.

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

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

Let Y_k be the k th 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).

¹¹ D is zero for all members of the control group. Models that include a treatment group and a control group are used to estimate π_1 while accounting for chance differences in X between treatment and control groups, in order to serve as the first stage of a 2SLS model in equation (3) below. Models for factors affecting take-up of the offered voucher in which β_1 is of direct interest are estimated separately within each treatment group and do not estimate π_1 .

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

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. Separate regression models are used to compare the experimental group to the control group and the Section 8 group to the control group.

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 an 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 approach relies on the assumption that there was no average effect of being offered an MTO voucher on those who did not use an MTO voucher, which we believe is a reasonable approximation, but not strictly true.¹³

Using terminology of Joshua D. Angrist, Guido W. Imbens, and Donald B. Rubin (1996), we refer to those using an MTO voucher to move as treatment compliers. The treatment

¹² Inflation by the proportion in the treatment group who actually received the treatment was introduced in the program evaluation literature by Howard Bloom (1984); see James Heckman, Robert LaLonde, and Jeffrey Smith (1999) for a comprehensive discussion of alternative parameters of interest in the evaluation of social programs.

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.

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 (Lawrence F. Katz, Jeffrey R. Kling, and Jeffrey B. Liebman 2001). CCM is defined in equation (4) for each outcome k .

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

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. The building blocks of the summary measures are standardized treatment effect sizes. Let σ_k^2 equal the variance of Y_k for the control group.¹⁴ 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.¹⁵

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

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

¹⁴ We use the control group standard deviation to compare the treatment groups to their counterfactual. This metric does not depend on which treatment (experimental or Section 8) is being analyzed.

¹⁵ If there were no missing data on survey items and X contains only a constant, then this summary measure is 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 6 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.

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).

$$(6) \quad Y = (I_K \otimes (Z \ X))\theta + v \quad Y = (Y_1', \dots, Y_K)'$$

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).

C. Data

The data in this study come from baseline surveys administered prior to random assignment, from administrative data, and from the MTO interim impacts survey conducted in 2002. The administrative data on earnings are from the Unemployment Insurance (UI) and state and county welfare agencies of California, Illinois, Maryland, Massachusetts, and New York.¹⁶ The analysis of administrative data reported here was conducted in collaboration with Abt Associates in preparation for our report (Larry 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.¹⁷ The survey had an effective response rate of 90 percent.¹⁸ Response rates were 91 percent for the experimental group and 89 percent for both

¹⁶ 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 for that state.

¹⁷ 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.

¹⁸ 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.

the Section 8 and control groups. All statistical estimates in this paper use survey weights.¹⁹

The baseline covariates (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

Overall, this sample of MTO adults is mainly female and black or Hispanic, as shown in Table 1.²⁰ 93 percent were ages 25-54 as of December 31, 2001. At the time of random assignment, three quarters were on AFDC, more than half had never married, less than half had graduated from high school, and a quarter had been teenage parents. There were two statistically significant differences (at the 5 percent level) between the two treatment groups and the control group in 72 comparisons using the covariates in Table 1.²¹

The fraction using a voucher offered by MTO to move -- which we refer to as the compliance rate -- was 47 percent for the experimental group and 60 percent for the Section 8 group.²² 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

¹⁹ 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.

²⁰ For 85 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.

²¹ 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 five polynomials 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.

²² These compliance rates are calculated using the sample weights and, therefore, differ slightly from the raw sample counts shown in Table 1.

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.²³ 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.²⁴ 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.²⁵

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.²⁶ Annual poverty rates for a census tract are obtained by linear interpolation from the rates estimated in the 1990 and 2000 censuses. 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.

²³ 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.

²⁴ For this analysis, equation (1) is estimated separately for the experimental group only and the Section 8 group only. There is no compliance in the control group.

²⁵ 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. The tightness of local housing markets presumably mattered as well.

²⁶ 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 differences between the tract poverty rates for treatment compliers and the implied distribution for control compliers are starker.²⁷ 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 2.²⁸ About 13 percent of the control complier density lies below 30 percent poverty in both the experimental and Section 8 groups, which 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.²⁹

Further contextual factors are shown in Table 2 -- selected to be illustrative, with a more complete set examined in section III.F. In addition to having lower poverty rates, the census tracts where treatment group members resided 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 groups 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

²⁷ Although initial moves through MTO in the experimental group were restricted to census tracts with poverty rates less than ten percent, this restriction lasted for only one year. In addition, these initial destination tracts had higher poverty rates in 2000 than in 1990.

²⁸ 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. Guido W. Imbens and Donald B. 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.

²⁹ 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.

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 controls.

The treatment had relatively modest impacts on the average census tract percentage minority, and few treatment group compliers moved to a tract that was less than 50 percent minority or moved more than 10 miles from their origin location. Relative to the control group, the housing quality of the units the experimental group moved to was somewhat higher in terms of problems with vermin, while the difference for the Section 8 group was negligible. Measures of social networks similarly show modest impacts overall. For example, 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. There were also no significant differences among the experimental groups in the frequency of church attendance.

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 initial 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 contextual factors.³⁰ 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 a greater fraction of working adults would provide MTO movers with stronger networks for job seeking, stronger community norms supportive of work, and physical proximity that would facilitate job search and possibly reduce commuting. 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.³¹

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, 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 both employed and not on TANF than the control group. None of these treatment effects are statistically significant,

³⁰ Results for additional outcomes, pre-specified in the study design as less important than the ones presented in the main results, are available in a web appendix as described in Appendix Table A0.

³¹ Bruce A. Weinberg, Patricia B. Reagan, and Jeffrey J. Yankow (2004) analyzed how work hours are related to neighborhoods in the NLSY using individual fixed effects. Susan J. Popkin, James E. Rosenbaum, and Patricia M. Meaden (1993) analyzed suburban and central city residents in the Gautreaux program using survey data, and Ruby Mendenhall, Greg J. Duncan, and Stephanie DeLuca (2003) followed-up longer-term outcomes with administrative data. Paul 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 (2004); Katz, Kling, and Liebman (2001); Hanratty, McLanahan, Becky Petit (2003); and Leventhal and Brooks-Gunn (2003b).

although we note that we do not have the statistical power to detect changes less than 10 percent of the control group mean.³² 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 wellbeing than to self-sufficiency per se, such as total income and food security, also show no statistically significant effects.

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 2002.³³ There was a substantial increase in incentives for work from welfare reform and the Earned Income Tax Credit, and welfare rolls decreased dramatically and employment rates rose for female heads (Rebecca Blank and David T. Ellwood, 2002). The decline in welfare income and increase in earnings had roughly offsetting effects on the consumption of single mothers in the late 1990s (Bruce D. Meyer and James X. Sullivan, 2004). In addition, as time passed, the adults in our sample were less likely to have small children at home. As shown in Figure 3, the MTO population experienced large changes over this period, with employment rates rising from less than 25 percent in early 1995 to more than 50 percent in 2001.³⁴ In 2002, the occupations of

³² 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 (Larry 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.

³³ 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).

³⁴ For our analyses based on UI data, “employment” is defined as positive earnings in a quarter and no UI record of earnings in the quarter is imputed to be 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. When all earnings for some individuals in a quarter were outside the state of random assignment, then the administrative data employment rate will be too low. 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 the full sample.

MTO sample adults were mainly in white-collar and service occupations, with hourly wages ranging from \$7.44 at the 25th percentile to \$12.47 at the 75th percentile.³⁵

The lack of differences between the MTO groups in administrative data on employment in Figure 3 is reflected in the coefficients in panel A of Table 4 which shows employment results for 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).³⁶ The results on earnings for these time periods in panel B and for public assistance receipt in panel C similarly fail to exhibit any statistically significant effects.³⁷

We do find some evidence that the neighborhoods the experimental group moved into may have been experiencing job loss instead of the job growth that we had hypothesized would occur. The experimental group moved into neighborhoods where the level of employment was substantially higher but where the trend in employment among residents was actually declining a bit faster relative to the control group.³⁸ If the change in employment among residents reflects

³⁵ The data on occupation and hourly wage are taken from survey self-reports about the main job in the week of 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.

³⁶ 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. 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) and the cell data for MA gives very similar point estimates with standard errors that are about five percent smaller.

³⁷ We do not find effects on welfare in 2001 or five years after RA 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 RA found by Ludwig, Duncan, and Pinkston (2004) did not persist.

³⁸ Nationally, the employment to population (e/pop) ratio for ages 16+ was .653 in 1990 and .639 in 2000, for a decline of .014 (Sandra L. Clark and Mai Weismantle, 2003). In results not shown in the table, the current census tract locations (at the time of the survey in 2002 for the control group had an average 1990 e/pop ratio of .433 and a 2000 ratio of .412, for a decline of .021. The current locations of the experimental group had a 1990 e/pop ratio of

the hiring patterns of accessible employers, then there may not have been more job vacancies available to the experimental group. We also find some evidence that the social networks relevant to employment for this population may not have been affected by 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.³⁹ We found no significant between group differences on commuting-related outcomes.⁴⁰

The consistency between the estimates we obtain using survey and administrative measures of economic self-sufficiency increases our confidence that differences between MTO groups in self-reported measures are meaningful, more generally. In planning for the study, it was suggested to us that our survey responses would give the appearance of more beneficial outcomes for groups who were offered vouchers because respondents would “tell us what they thought we wanted to hear,” but that these effects would not be verified using administrative data.⁴¹ Given that the name of the demonstration is “Moving to Opportunity” and that it was promoted by the government 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. Other work on youth arrest records in MTO also finds no evidence of

.499 and a 2000 ratio of .466, for a decline of .033. The difference-in-difference of $-.033 - (-.021) = -.012$ has a p-value of .002. Of course, this pattern could be driven by changes in the composition of residents in these particular tracts and not by job loss in the area. The declining trajectory of the neighborhoods to which the experimental group moved was also evident in the differences in tract poverty for given locations as measured in 1990 versus 2000. The current locations of the experimental group had an average increase in tract poverty rate from 1990 to 2000 of .006, while Section 8 group had virtually no change and the control group had a decrease in tract poverty of .011.

³⁹ The types of informal networks emphasized by Patrick Bayer, Stephen L. Ross, and Giorgio Topa (2004) may exist for this sample, but may not be of large importance.

⁴⁰ 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.”

⁴¹ It was also suggested that interviewers themselves, though not told whether the person was a member of the intervention group, might surmise which group they were in from where the person lives and somehow administer the questions or record the answers differently.

respondents in households offered vouchers reporting more socially desirable outcomes (Kling, Ludwig, and Katz, 2004).

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.⁴² 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

⁴² 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 non-specific 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.

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.⁴³ In the MTO sample, the control complier mean for the fraction with SMI is .17, and there is a TOT effect of -.076 (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). These results are consistent with significant reductions in distress within the more-distressed half of the complier population, under an assumption of perfect rank correlation between treatment and control complier outcomes.

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 marginally significant TOT effect of a .058 reduction in probability of a DSM-IV diagnosis of a major depressive episode (p-value=.06; CCM=.194), 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.⁴⁴ The effects on worrying

⁴³ SMI is defined as having a non-substance-use DSM-IV 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 Ronald C. Kessler et al (2003) for details.

⁴⁴ For example, in a study often cited as an exemplar of an effective intervention, Kenneth B. Wells et al. (2000) analyzed outcomes of depressive patients randomized to obtain usual care or improved quality care (better training

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 3526 adults observed on average 5 years after random assignment are broadly consistent with the early results for two of the MTO sites.⁴⁵

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.⁴⁶

C. Physical Health

We hypothesized that moving to new residential locations could decrease environmental irritants, provide safer locations for exercise and generally less stress, or increase access to more

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.

⁴⁵ 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, Tama Leventhal and Jeanne 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.

⁴⁶ The largest estimates of the effect of mental health on employment that we are aware of are those in Susan L. Ettner, Richard G. Frank, and Ronald C. Kessler (1997) who report, for example, that a current episode of major depression reduces the odds of female employment by about 38 percent. Starting from our control mean level of employment of 55.5 percent, an intervention that converted everyone in our sample from depressed to non-depressed would therefore increase employment to 66.7 percent. Our estimated 2.7 percentage point reduction in experimental group depression would therefore translate into a 0.30 percentage point increase in employment $(66.7 - 55.5) \times .027$ starting from the control mean of 55.5 percent. Even if the intervention had completely eliminated depression among our sample, we would only expect to see a 1.83 percentage point increase in employment, well below our minimum detectable effect. These estimates implausibly assume that the effect of mental health on employment operates only through crossing the threshold from non-depressed to depressed. Allowing for employment effects through improvements in mental health that did not result in changes in the depression classification would raise the expected employment effects, but it seems quite unlikely that they would raise them by enough to alter the conclusion that such effects are well below our minimum detectable effects.

nutritious foods. Peers in lower poverty neighborhoods may have healthier eating and exercise habits. Health care access could also change. Effects on such contextual 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 $BMI \geq 30 \text{ kg/m}^2$, we find that nearly half of the control group would be classified as obese.⁴⁷ 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).

D. *Summary measures of main outcomes*

In brief, we find that aggregating the information for specific outcomes in the domains of economic self-sufficiency, mental health, and physical health amplifies our conclusions above.

The results for mean effect sizes using the full sample, shown in the first two columns of Table

⁴⁷ 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 $BMI \geq 30$ (Jacqueline W. Lucas, Jeannine S. Schiller, and Veronica Benson 2004). Our data on weight is 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 (Janet L. 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 (Elmer V. Villanueva 2001). Changes in neighborhood could affect reporting of weight in two offsetting ways. First, movers might perceive themselves as heavier than those around them and over-report their weight. Alternatively, movers might feel more self-conscious about their weight in new neighborhoods and under-report.

6, are based on the equally weighted standardized treatment effects described in equation (5).

We interpret these summary measure results for each domain as indices aggregating information about effects on related constructs, but we do not intend to suggest that all measures within a domain are merely proxies for a single latent factor.⁴⁸ We do not observe significant mean effects on any domain for the Section 8 group, and we do not observe significant mean effects on economic self-sufficiency or physical health for the experimental group. There is a consistent and sizeable effect of being offered an experimental group voucher and moving to a lower poverty neighborhood on mental health that is unlikely to have occurred by chance.

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 .0052.⁴⁹ 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. 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: $1-(1-.0052)^8 = .041$. Using a modified

⁴⁸ 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.

⁴⁹ 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.

Westfall-Young method to account for dependence between the estimates described in detail by Kling and Liebman (2004), 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. For other domains and groups, the effects are positive in sign but statistically insignificant.

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.⁵⁰ The mean effect size estimates come from equation (6), where Z contains the treatment indicator 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 does have a pattern where greater reductions in poverty rates are associated with greater improvement in mental health. The pattern also 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. The pattern for mental health is contrary to a model in which the

⁵⁰ 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 Harry J. Holzer, Lawrence F. Katz, and Alan B. Krueger (1991). In related work (Jeffrey B. Liebman, Lawrence F. Katz, and Jeffrey R. Kling 2003), we have examined such models which essentially fit a line through the points in Figure 4. Note that the poverty rate is used as an index of neighborhood attributes, and is not interpreted as the effect of a change in the poverty rate holding other neighborhood attributes constant.

unrestricted choice of the Section 8 group should have led to more beneficial outcomes than the restricted choice of the experimental group. Moreover, there are virtually no outcomes or contextual factors for which the Section 8 assignment was more beneficial than the experimental assignment, suggesting that there was unlikely to have been offsetting value of Section 8 on some other dimension.⁵¹ 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 may well have led the restriction to in the end be beneficial for the individuals on whom it was imposed.

E. *Summary measures of outcomes for subgroups*

We have also explored whether the effects differ by baseline characteristics, such as race and education. In general, results do not differ appreciably by these characteristics.⁵² Similarly, we find that our main conclusion (mean 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 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, and we note this type of searching for significant effects in subgroups will in general result in a greater chance of the appearance of statistical significance even when the null hypothesis of no effects is true.

⁵¹ The experimental group did have a lower compliance rate than the Section 8 group, and it is possible that individuals who would have complied under both treatment offers actually had equivalent effects or even more beneficial effects for Section 8, but the effects were smaller for those who would have only complied with the Section 8 offer.

⁵² 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 the beneficial experimental group mental health effects 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.

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 8.⁵³ 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 susceptible to a change in environmental circumstances.

A similar pattern of relatively more beneficial results for younger adults holds for economic self-sufficiency in the experimental group at a lower level of significance (p-value = .13). 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.⁵⁴ This pattern of results for the younger

⁵³ The estimates for the 15 individual outcomes are shown in Appendix Table A1.

⁵⁴ 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 \leq 32 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.

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.

F. *Summary measures of contextual 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 contextual factors, using the same mean effect size metric as in Table 6.⁵⁵ Our intention in collecting extensive data on contextual factors was to explore mechanisms through which changes in residential location might impact economic and health outcomes.

Relative to the standard deviation in the control group, the changes in tract characteristics (poverty rate, female employment rate, 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, consistent with the effects on obesity for those groups.

⁵⁵ Results for each component of the summary measures of contextual factors are given in the web appendix described in Appendix Table A0.

Given the large differences in physical health outcomes by age found in Table 6, we suspected there might be differences by age in health care access or exercise and nutrition, but these were not observed in the data. Even the differences in effects by age on smoking and drinking (with older adults in the experimental group smoking and drinking more than controls) are 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. Assuming no effects for those who did not use a voucher to move, the differences in mean effect sizes by age for the treated on tract characteristics and on safety are not significant. 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 our measures of neighborhood or social context.

G. Internal and External Validity

This study has several limitations. First, 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.⁵⁶

Second, our effective survey response rate was 90 percent and it is possible that characteristics of those who were not interviewed differed systematically across the three groups.

⁵⁶ Additional supporting evidence comes from related work (Kling and Liebman 2004) which finds strong beneficial effects on the mental health of female youth. 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.

Our estimation models control for baseline characteristics in order to reduce the sensitivity of our results to differential attrition. Of course, it is also possible that the individuals who were not interviewed in the three groups differed in their unobservable characteristics.⁵⁷ 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.⁵⁸

Third, about half of the subjects randomized to the intervention groups actually used the voucher that was offered and some members of the control group later moved out of public housing, reducing our statistical power to detect effects of location on outcomes. Fourth, 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. Fifth, 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. Sixth, our results are most applicable to situations where the number of households moving into any particular neighborhood is small

⁵⁷ 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.

⁵⁸ 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. There was evidence that the subsampled individuals (pooling survey respondents and nonrespondents with no attrition) were different by the chance of the 3-in-10 subsampling. The p-values on the contrasts of pooled subsample vs. full sample ranged from .15 to .01.

(as it was in MTO), as larger numbers may have different “endogenous” or “general equilibrium” effects (Charles F. Manski 1993, James J. Heckman, 2001).

IV. Conclusion

Broadly speaking, we use an experiment to address the issue of self-selected residential location and find evidence for neighborhood effects on some but not all outcomes. A randomized intervention that used housing vouchers to encourage families in high-poverty neighborhoods to move to low-poverty neighborhoods was associated with better mental health among adults in an experimental group relative to a control group. However, overall there were no discernable improvements in economic self-sufficiency or physical health for the sample. We also observe no significant overall pattern of effects for self-sufficiency or health outcomes in the 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. The intervention also resulted in a relatively homogenous group of people living in a wide range of neighborhoods, with a magnitude of neighborhood variation rarely present in the non-experimental studies of neighborhood effects on low-income households. 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 health and reduced obesity is consistent with theories hypothesizing a link between the two conditions.⁵⁹

Relative to previous studies, 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.⁶⁰ 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.⁶¹ This study demonstrates that systematic variation in location induced by randomization can be accomplished, and such randomization overcomes many of the methodological shortcomings of the observational studies. Researchers need not wait for large federal demonstration programs to engage in this sort of research. For example, Philip Oreopoulos (2003) used quasi-random variation in location caused by a local housing authority in Canada to study the impact of neighborhoods on economic outcomes, and waiting lists and other bureaucratic allocation mechanisms could also be used to study effects of social environment on health.⁶²

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.

⁵⁹ Kenneth M. Carpenter, Deobrah S. Hasin, David B. Allison, and Myles S. Faith, 2000; Myles S. Faith, Patty E. Matz, and Marie A. Jorge, 2002; Michael A. Friedman and Kelly D. Brownell, 1995; Per Bjorntorp, 2001.

⁶⁰ Ellaway, Anderson, and Macintyre, 1997; Yen and Kaplan 1999b; Silver, Mulvey and Swanson 2002.

⁶¹ Ong, 1998; Robert, 1998; C. Kadushin et al, 1998; Craig Duncan, Kelwyn Jones, and Graham Moon, 1999; Ross and Mirowsky 2001.

⁶² Jacob (2004) provides an application of this approach to educational outcomes.

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 a 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.⁶³

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.⁶⁴ We hypothesize several possible reasons for the lack of an employment effect. First, employment rates were declining in the experimental group locations at rates even faster than those in the control group. Second the MTO intervention may not have caused enough people to move far enough for us to see effects in local labor market conditions even though tract poverty rates and the share of employed and middle-class neighbors changed dramatically. Third, there may have only been employment effects for certain subgroups or labor market conditions, such as younger adults in tighter labor markets, and not the more consistent overall effects we had originally hypothesized and on which we focused our main

⁶³ Susan J. Popkin, Laura E. Harris and Mary K. Cunningham, 2003; Kling, Liebman and Katz, 2004.

⁶⁴ One possibility is that spatial mismatch hypothesis is false – that when truly exogenous variation in residential location is available there is little or no impact of residential location on employment opportunities. Skeptics of the hypothesis include David T. Ellwood (1986) and Christopher Jencks and Susan E. Mayer (1990). It could also be argued that the spatial mismatch hypothesis was never a story about adult female workers, in that most of the spatial mismatch literature studies teenage workers and the original Wilson (1987) argument was primarily about male workers losing access to manufacturing jobs.

analysis. Fourth, we do not have the statistical precision to rule out moderate changes in employment and earnings in either direction being generated by the MTO treatments.

From a policy perspective, the overarching question about the role of government in providing housing assistance involves the effects on many groups of individuals including the adult participants studied here. These results for adults suggest that interventions that improve distressed neighborhoods or assist people in leaving them can have important public health benefits. The fact that there are no statistically significant overall impacts on earnings or welfare usage five years after random assignment suggests that mental health outcomes may be more sensitive than economic self-sufficiency outcomes to housing policies for low-income families, at least in the medium run. The mental health benefits may have important spillover benefits, particularly to children who have been found to have more problems in school and more behavior problems when their mothers are experiencing mental health problems.⁶⁵ Taken together, these findings 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.

⁶⁵ For reviews of linkage between maternal mental health and child development, see Cheryl T. Beck (1999) and Sherryl H. Goodman and Ian H. Gotlib (1999). Another possible spillover of improving mental health could be reduced Medicaid costs, although expenditures on mental health services are not large. Jeffrey A. 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.

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 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 DSM-IV psychiatric diagnosis. The probability is based on a mapping between the CIDI-SF screening questions and more detailed assessments in the National Comorbidity Survey (Ellen E. Walters et al 2003).⁶⁶

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 SF-36 (John E. Ware et al, 1993).

⁶⁶ Version 1.0 of the World Health Organization’s CIDI-SF contained a skip pattern error for people who reported dysphoric mood 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 for their symptoms of anhedonia. Assigning an MDE probability of zero to boundary cases is one solution to this problem. We implement an alternative imputation that is also conservative but less extreme; we impute the probability of anhedonia for boundary cases to be .26 using data observed for those with dysphoric mood less than half the day, under the assumption that the probability of anhedonia for cases with dysphoric mood for “about half of the day” would be at least this high. 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 .85, we then impute the probability of MDE for those receiving medication to also be .85. 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.

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 Tracy L. Skaer et al. (1999).

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 most 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 $\geq 140\text{mm HG}$ or diastolic $\geq 90\text{mm HG}$ (Chobanian, Aram V 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

- Anwar, Yusra Anis; Giacco, Sharon; McCabe, Ellen J.; Tandler, Beatriz E. and White, William B. "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*, 1998, 3, pp. 261-265.
- Bayer, Patrick; Ross, Stephen L. and Topa, Giorgio. "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes." Manuscript, Federal Reserve Bank of New York, 2004.
- Benabou, Roland. "Workings of a City: Location, Education, and Production." *Quarterly Journal of Economics*, 1993, 108:3, pp. 619-652.
- Beck, Cheryl H. "Maternal depression and child behavior problems: a meta-analysis." *Journal of Advanced Nursing*, 1999, 29(3), pp. 623-629.
- Bjorntorb, Per. "Do Stress Reactions Cause Abdominal Obesity and Comorbidities?" *Obesity Review*, 2001, 2, 73-86.
- Blank, Rebecca and David T. Ellwood. "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, 2002.
- Bloom, Howard. "Accounting for No-shows in Experimental Evaluation Designs." *Evaluation Review*, 1984, 8, pp. 225-246.
- Brock, William A. and Durlauf, Steven N. "Interactions-based Models." In James J. Heckman and Edward Leamer, eds, *Handbook of Econometrics, Volume 5*. Amsterdam: North-Holland, 2001, pp. 3297-3380.
- Buck, Jeffrey A.; Teich, Judith L.; Bae, Jay and Dilonardo, Joan. "Mental health and substance abuse services in ten state Medicaid programs." *Administration and Policy in Mental Health*. 2001, 28(3), 181-92.
- Carpenter, Kenneth M., Hasin, Deborah S., Allison, David B., and Faith, Myles S. "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*, 2000, 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. "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*, 2003, 289, pp. 2560-2572.
- Clark, Sandra L. and Weismantle, Mai. *Employment Status: 2000*. Washington, DC: U.S. Census Bureau, 2003.
- Committee on Ways and Means. *The 2000 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, 2000.
- Committee on Ways and Means. *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, 2004.
- Cutler, David M. and Glaeser, Edward L. "Are Ghettos Good or Bad?" *Quarterly Journal of Economics*, 1997, 112(3), 827-872.

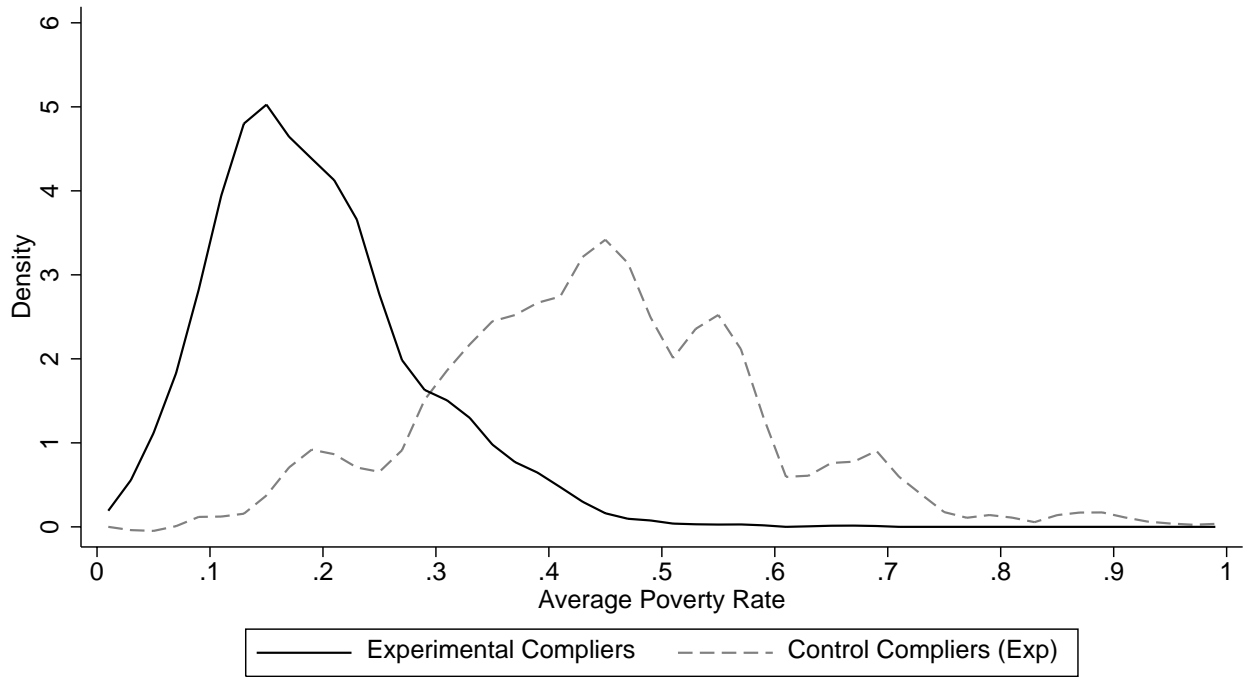
- Diez-Roux, Ana V.; Merkin, Sharon Stein; Arnett, Donna; Chambless, Lloyd; Massing, Mark; Nieto, F. Javier; Sorlie, Paul; Szklo, Moyses; Tyroler, Herman A. and Watson, Robert L. "Neighborhood of residence and incidence of coronary heart disease." *New England Journal of Medicine*, 2001, 345, pp. 99-106.
- Duncan, Craig; Jones, Kelwyn and Moon, Graham. "Smoking and Deprivation: Are there neighbourhood effects?" *Social Science and Medicine*, 1999, 48, 497-505.
- Ellaway, A.; Andersen, A. and Macintyre, S. "Does Area of Residence Affect Body Size and Shape?" *International Journal of Obesity*, 1997, 21, pp. 304-308.
- Ellwood, David T. "The Spatial Mismatch Hypothesis: Are There Teenage Jobs Missing In the Ghetto?" In Richard B. Freeman and Harry J. Holzer, eds., *The Black Youth Employment Crisis*. Chicago: University of Chicago Press, 1986, pp. 147-187.
- Engstrom, Janet L.; Paterson, Susan A.; Doherty, Anastasia; Trabulsi, Mary and Speer, Kara L. "Accuracy of Self-reported Height and Weight in Women: An Integrative Review of the Literature." *Journal of Midwifery & Women's Health*, 2003, 48(5), 338-345.
- Ettner, Susan L., Richard G. Frank, and Ronald C. Kessler. "The Impact of Psychiatric Disorders on Labor Market Outcomes." *Industrial and Labor Relations Review*, 1997, 51, 64-81.
- Faith, Myles, S., Matz, Patty E., and Jorge, Marie A. "Obesity-Depression Associations in the Population." *Journal of Psychosomatic Research*, 2002, 52, 935-942.
- Feins, Judith D.; McInnis, Debra and Popkin, Susan J. *Counseling in the Moving to Opportunity Demonstration Program*. Washington, DC: U.S. Department of Housing and Urban Development, 1997.
- Friedman, Michael A. and Brownell, Kelly D. "Psychological Correlates of Obesity: Moving to the Next Research Generation." *Psychological Bulletin*, 1995, 117, 3-20.
- Furukawa, Toshiaki A.; Kessler, Ronald C.; Slade, T. and Andrews, G. "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*, 2003, 33, 357-362.
- Goodman, Sherryl H. and Gotlib, Ian H. "Risk for Psychopathology in the Children of Depressed Mothers: A Developmental Model for Understanding Mechanisms of Transition." *Psychological Review*, 1999, 10(3), pp. 458-490.
- Granovetter, Mark. *Getting a Job: A Study of Contacts and Careers*. Cambridge, MA: Harvard University Press, 1974.
- Heckman, James J. "Accounting for Heterogeneity, Diversity and General Equilibrium in Evaluating Social Programs." *Economic Journal*, 2001, 111, pp. 654-699.
- Heckman, James J.; LaLonde, Robert J. and Smith, Jeffrey A. "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, 1999.
- Hanratty, Maria; McLanahan, Sara and Pettit, Becky. "Los Angeles Site Findings," in John Goering and Judith D. Feins, eds., *Choosing a Better Life: Evaluating the Moving to Opportunity Social Experiment*. Washington, DC: The Urban Institute Press, 2003, pp. 245-274.
- Holzer, Harry J.; Katz, Lawrence F. and Krueger, Alan B. "Job Queues and Wages." *Quarterly Journal of Economics*, 1991, 106(3), pp. 739-768.
- Ihlanfeldt, Keith R. and Sjoquist, David J. "The Spatial Mismatch Hypothesis: A Review of Recent Studies and Their Implications for Welfare Reform." *Housing Policy Debate*, 1998, 9(4), pp. 849-892.

- Imbens, Guido W. and Rubin, Donald B. "Estimating outcome distributions for compliers in instrumental variables models." *Review of Economic Studies*, 1997, 64(221), pp. 555-575.
- Jacob, Brian A. "Public Housing, Housing Vouchers, and Student Achievement: Evidence from Public Housing Demolitions in Chicago." *American Economic Review*, 2004, 94(1), pp. 233-258.
- Jargowsky, Paul A. *Poverty and Place*. New York: Russell Sage, 1997.
- Jencks, Christopher, and Mayer, Susan E. "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, 1990, pp. 187-222.
- Kadushin C.; Reber, E.; Saxe, L. and Livert, D. "The substance use system: social and neighborhood environments associated with substance use and misuse." *Substance Use and Misuse*, 1998, 33, pp. 1681-1710.
- Kasarda, John D. "Urban Industrial Transition and the Underclass." *Annals Of The American Academy Of Political And Social Science*, 1989, 501(1), pp. 26-47.
- Kain, John F. "Housing Segregation, Negro Employment and Metropolitan Decentralization." *Quarterly Journal of Economics*, 1968, 82, pp. 175-197.
- Katz, Lawrence F.; Kling, Jeffrey R. and Liebman, Jeffrey B. "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment." *Quarterly Journal of Economics*, 2001, 116(2), pp. 607-654.
- Kawachi, Ichiro and Berkman, Lisa F. "Introduction," in Ichiro Kawachi and Lisa F. Berkman, eds., *Neighborhoods and Health*. Oxford: Oxford University Press, 2003, pp. 1-19.
- Kessler, Ronald C.; Andrews G.; Mroczek, D.K.; Ustun. T.B. and Wittchen, H.U. "The World Health Organization Composite International Diagnostic Interview Short-Form (CIDI-SF)." *International Journal of Methods in Psychiatric Research*, 1998, 7, pp. 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. "Short screening scales to monitor population prevalences and trends in non-specific psychological distress." *Psychological Medicine*, 2002, 32, pp. 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. "Screening for serious mental illness in the general population." *Archives of General Psychiatry*, 2003, 60, pp. 184-189.
- Kling, Jeffrey R. and Liebman, Jeffrey B. "Experimental Analysis of Neighborhood Effects on Youth." Princeton IRS Working Paper 483, 2004.
- Kling, Jeffrey R.; Liebman, Jeffrey B. and Katz, Lawrence F. "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, 2004.
- Kling, Jeffrey R.; Ludwig, Jens and Katz, Lawrence F. "Youth Criminal Behavior in the Moving To Opportunity Experiment." Princeton IRS Working Paper 482, 2004.
- Leventhal, Tama and Brooks-Gunn, Jeanne. "Moving To Opportunity: an Experimental Study on Neighborhood Effects on Mental Health." *American Journal of Public Health*, 2003a, 93(9), pp. 1576-1582.
- Leventhal, Tama and Brooks-Gunn, Jeanne. "New York City Site Findings: The Early Impacts of Moving To Opportunity on Children and Youth," in John Goering and Judith D. Feins,

- eds., *Choosing a Better Life: Evaluating the Moving to Opportunity Social Experiment*. Washington, DC: The Urban Institute Press, 2003b, pp. 213-244.
- Liebman, Jeffrey B.; Katz, Lawrence F., and Kling, Jeffrey R. "Are Neighborhood Effects Nonlinear? Evidence from the Moving to Opportunity Experiment." Unpublished paper, NBER, 2003.
- Lucas, Jacqueline W.; Schiller, Jeannine S. and Benson, Veronica. *Vital and Health Statistics, Series 10, Number 218*. Washington, DC: National Center for Health Statistics, 2004.
- Ludwig, Jens; Duncan Greg J. and Pinkston, Joshua C. "Housing mobility programs and economic self-sufficiency: Evidence from a randomized experiment." *Journal of Public Economics*, 2004, forthcoming.
- Manski, Charles F. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies*, 1993, 60, pp. 531-542.
- Mendenhall, Ruby; Duncan, Greg J. and DeLuca, Stephanie. "Neighborhood Resources and Economic Mobility: Results from the Gautreaux Experiment." Manuscript, Northwestern University, 2003.
- Meyer, Bruce D. and Sullivan, James X. "The effects of welfare and tax reform: the material wellbeing of single mothers in the 1980s and 1990s." *Journal of Public Economics*, 2004, 88, pp. 1387-1420.
- Montgomery, James D. "Social Networks and Labor Market Outcomes: Toward an Economic Analysis." *American Economic Review*, 1991, 81(5), pp. 1408-1418.
- Montgomery, James D. "Job Search and Network Composition: Implications of the Strength-of-Weak-Ties Hypothesis." *American Sociological Review*, 1992, 57(5), pp. 586-596.
- National Institutes of Health. "Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults--The Evidence Report." *Obesity Research*, 1998, 6(Supp 2), pp. 51S-209S.
- Olsen, Edgar O. "The Cost-Effectiveness of Alternative Methods of Delivering Housing Subsidies." Manuscript, University of Virginia, 2000.
- Olsen, Edgar O. "Housing Programs for Low-Income Households," in Robert A. Moffitt, ed., *Means-Tested Transfer Programs in the United States*. Chicago: University of Chicago Press and NBER, 2003, pp. 365-441.
- Ong, Paul. "Subsidized Housing and Work among Welfare Recipients." *Housing Policy Debate*, 1998, 9(4), pp. 775-794.
- Orr, Larry. *Social Experiments*. London: Sage Publications, 1999.
- Orr, Larry; Feins, Judith D.; Jacob, Robin; Beecroft, Eric; Sanbonmatsu, Lisa; Katz, Lawrence F.; Liebman, Jeffrey B. and Kling, Jeffrey R. *Moving to Opportunity: Interim Impacts Evaluation*. Washington D.C.: U.S. Department of Housing and Urban Development, 2003.
- Oreopoulos, Philip. "The Long-run Consequences of Growing Up in a Poor Neighborhood," *Quarterly Journal of Economics*, 2003, 118(4), pp. 1533-1575.
- Popkin, Susan J.; Harris, Laura E. and Cunningham, Mary K. *Families in Transition: A Qualitative Analysis of the MTO Experience*. Washington, DC: Urban Institute, 2001.
- Popkin, Susan J.; Rosenbaum, James E. and Meaden, Patricia M. "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*, 1993, 12:3, pp. 556-573.
- Quigley, John M. "A Decent Home: Housing Policy in Perspective." *Brookings Papers on Urban Affairs*, 2000, 1, 53-99.

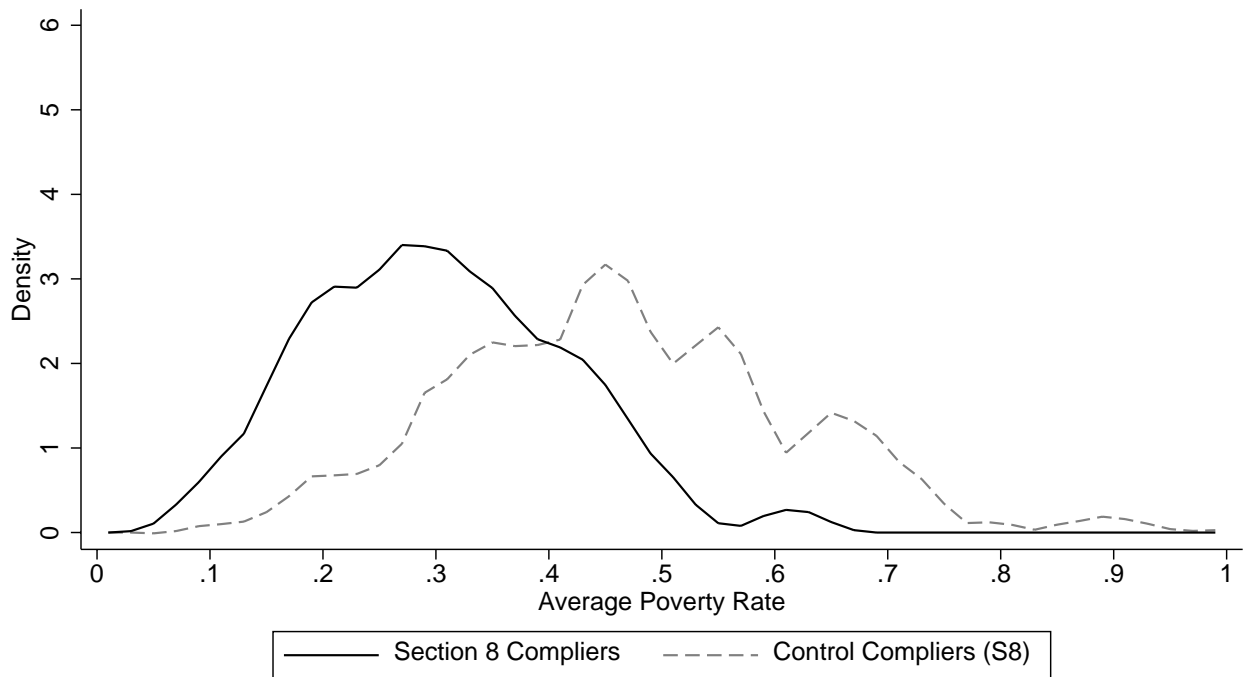
- Robert, Stephanie A. "Community-level socioeconomic status effects on adult health." *Journal of Health and Social Behavior*, 1998, 39, pp. 18-37.
- Ross, Catherine E. and Mirowsky, John. "Neighborhood disadvantage, disorder, and health." *Journal of Health and Social Behavior*, 2001, 42, pp. 258-276.
- Silver, Eric; Mulvey, Edward P. and Swanson, Jeffrey W. "Neighborhood structural characteristics and mental disorder: Faris and Dunham revisited." *Social Science & Medicine*, 2002, 55, pp. 1457-1470.
- Skaer, Tracy L.; Robison, Linda M.; Sclar, David A. and Galin, Richard S. "Psychiatric Comorbidity and Pharmacological Treatment Patterns among Patients Presenting with Insomnia." *Clinical Drug Investigation*, 1999, 18(2), pp. 161-168.
- Topa, Giorgio. "Social Interactions, Local Spillovers, and Unemployment." *Review of Economic Studies*, 2001, 68, 261-295.
- U.S. Department of Housing and Urban Development. "Economic Cost Analysis of Different Forms of Assisted Housing." Office of Policy Research and Development Issue Brief II, December. Washington, DC, 2000.
- Villanueva, Elmer V. "The Validity of Self-reported Weight in U.S. Adults: A Population Based Cross-sectional Study." *BMC Public Health*, 2001, 1:11.
<http://www.biomedcentral.com/1471-2458/1/11>
- Waltzman, Norman J. and Smith, Ken R. "Phantom of the area: poverty-area residence and mortality in the United States." *American Journal of Public Health*, 1998, 88, pp. 973-976.
- Walters Ellen E.; Kessler, Ronald C.; Nelson Christopher B. and Mroczek, Daniel M. "Scoring the World Health Organization's Composite International Diagnostic Interview Short Form." <http://www3.who.int/cidi/CIDISFScoringMemo12-03-02.pdf>. Accessed July 3, 2003.
- Ware, John E.; Snow, K.K.; Kosinski, M. and Gandek, B. *SF-36 Health Survey: Manual and Interpretation Guide*. Boston: The Health Institute, New England Medical Center, 1993.
- Weinberg, Bruce A.; Reagan, Patricia B. and Yankow, Jeffrey J. "Do Neighborhoods Affect Hours Worked: Evidence from Longitudinal Data." *Journal of Labor Economics*, 2004, forthcoming.
- Wells, Kenneth B.; Sherbourne, Cathy; Schoenbaum, Michael; Duan, Naihua; Meredith, Lisa; Unützer, Jürgen; Miranda, Jeanne; Carney, Maureen F. and Rubenstein, Lisa V. "Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial." *Journal of the American Medical Association*, 2000, 283, pp. 212-220.
- Wilner, Daniel M.; Walkley, Rosabelle P.; Pinkerton, Thomas C. and Tayback, Matthew. *The Housing Environment and Family Life: A Longitudinal Study of the Effects for Housing on Morbidity and Mental Health*. Baltimore: Johns Hopkins Press, 1962.
- Wilson, William J. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press, 1987.
- Wilson, William Julius. *When Work Disappears: The World of the New Urban Poor*. New York: Alfred A. Knopf, 1996.
- Yen, Irene H. and Kaplan, George A. "Neighborhood social environment and risk of death: multilevel evidence from the Alameda County Study." *American Journal of Epidemiology*, 1999a, 149, pp. 898-907.
- Yen, Irene H. and Kaplan, George A. "Poverty area residence and changes in depression and perceived health status: evidence from the Alameda County Study." *International Journal of Epidemiology*, 1999b, 28, pp. 90-94.

FIGURE 1. EXPERIMENTAL GROUP POVERTY RATE DENSITY



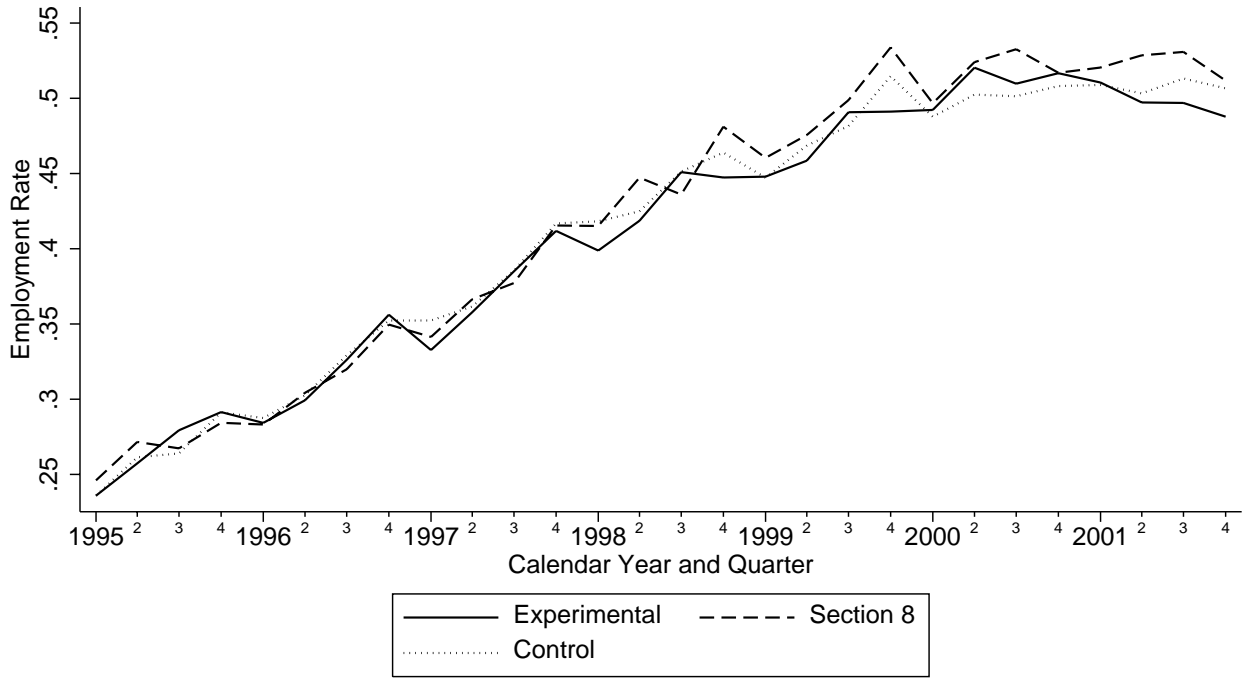
Notes. Kernel density based on actual distribution for experimental compliers and implied density for control compliers, as described in the text. Poverty rate averaged over tracts since random assignment, weighted by duration.

FIGURE 2. SECTION 8 GROUP POVERTY RATE DENSITY



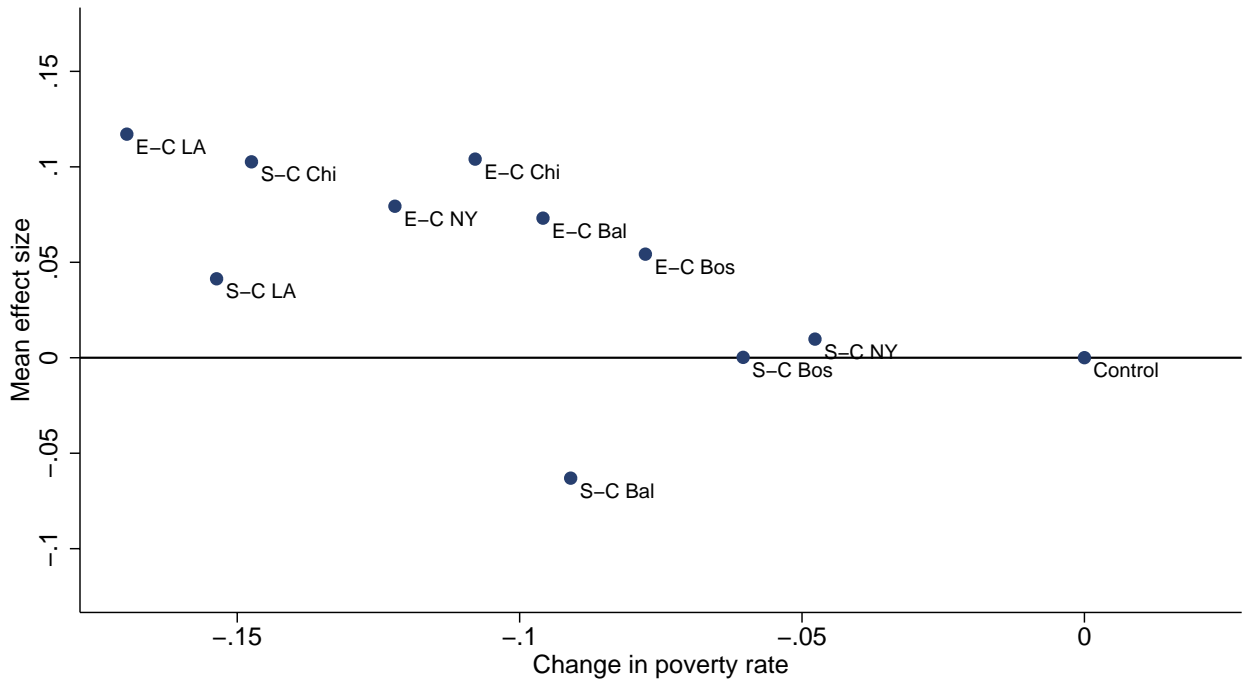
Notes. Kernel density based on actual distribution for Section 8 compliers and implied density for control compliers, as described in the text. Poverty rate averaged over tracts since random assignment, weighted by duration.

FIGURE 3. EMPLOYMENT RATES OVER TIME



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 (i)	Mean (ii)	CP Mean (iii)	NCP Mean (iv)	CP- NCP (v)	Mean (vi)	CP Mean (vii)	NCP Mean (viii)	CP- NCP (ix)
<u>Demographics</u>									
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*
<u>Economic and Education</u>									
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 General equivalency diploma	.38	.41	.41	.42	-.01	.41	.41	.40	.01
	.21	.18	.21	.15	.06*	.19	.20	.18	.01
<u>Household</u>									
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
<u>Neighborhood and Housing</u>									
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*

Notes: 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 1 — BASELINE CHARACTERISTICS, CONTINUED

Variable	Control	Experimental			Section 8				
	Mean (i)	Mean (ii)	CP Mean (iii)	NC Mean (iv)	CP- NCP (v)	Mean (vi)	CP Mean (vii)	NCP Mean (viii)	CP- NCP (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: 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 CONTEXTUAL FACTORS

	Experimental versus Control					Section 8 versus Control			
	CM (i)	ITT (ii)	TOT (iii)	CCM (iv)	N (v)	ITT (vi)	TOT (vii)	CCM (viii)	N (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 share of adults employed [ADDR]	.383	.074* (.004)	.159* (.008)	.385	2533	.056* (.004)	.093* (.006)	.378	2073
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 is 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. 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 (i)	Experimental versus Control				Section 8 versus Control			
		ITT (ii)	TOT (iii)	CCM (iv)	N (v)	ITT (vi)	TOT (vii)	CCM (viii)	N (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. 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 (i)	Experimental versus Control				Section 8 versus Control			
		ITT (ii)	TOT (iii)	CCM (iv)	N (v)	ITT (vi)	TOT (vii)	CCM (viii)	N (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 and food stamps									
Fraction of quarters received TANF in year 5 after RA [ADMIN]	.266	-.009 (.019)	-.020 (.041)	.287	1847	.018 (.022)	.029 (.036)	.260	1447
Amount of TANF payments received in year 5 after RA [ADMIN]	1295	-7 (105)	-15 (233)	1418	1847	10 (120)	16 (194)	1256	1447

Notes. ADMIN = administrative data. RA = random assignment. 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

	CM (i)	Experimental versus Control				Section 8 versus Control			
		ITT (ii)	TOT (iii)	CCM (iv)	N (v)	ITT (vi)	TOT (vii)	CCM (viii)	N (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 depression (lower bound) [SR]	.163	-.027 (.014)	-.058 (.031)	.194	2529	-.013 (.016)	-.021 (.027)	.163	2070
Worrying [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 \geq 30 [SR]	.468	-.048* (.022)	-.103* (.047)	.502	2450	-.046 (.025)	-.077 (.041)	.491	1999
Has hypertension, [M] SBP \geq 140 or DBP \geq 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

	Overall		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 [5 measures]	.016 (.031)	.034 (.034)	.063 (.046)	.047 (.049)	-.032 (.043)	.020 (.048)	.095 (.063)	.027 (.068)
Mental health [5 measures]	.084* (.030)	.030 (.034)	.076 (.043)	.028 (.047)	.093* (.042)	.031 (.048)	-.017 (.060)	-.002 (.066)
Physical health [5 measures]	.016 (.024)	.017 (.027)	.072* (.035)	.080* (.038)	-.035 (.035)	-.041 (.039)	.107* (.049)	.121* (.054)
Overall [15 measures]	.039 (.020)	.027 (.022)	.070* (.029)	.052 (.030)	.008 (.029)	.003 (.033)	.062 (.041)	.049 (.045)

Notes. RA = random assignment. 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 (iv) 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 CONTEXTUAL FACTORS

	Overall		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]	.666* (.034)	.484* (.033)	.757* (.049)	.537* (.046)	.572* (.049)	.431* (.048)	.184* (.069)	.106 (.066)
Neighborhood condition [3 meas.]	.253* (.034)	.192* (.036)	.276* (.047)	.201* (.051)	.231* (.049)	.182* (.052)	.044 (.067)	.019 (.073)
Safety [3 measures]	.220* (.029)	.187* (.030)	.277* (.040)	.253* (.043)	.161* (.042)	.118* (.045)	.116* (.058)	.135* (.062)
Housing quality [4 measures]	.133* (.029)	.096* (.032)	.167* (.041)	.109* (.046)	.097* (.040)	.081 (.044)	.070 (.058)	.028 (.063)
Social networks [11 measures]	.007 (.016)	.022 (.018)	.012 (.023)	.027 (.026)	.003 (.022)	.019 (.025)	.009 (.032)	.008 (.036)
Education and training [2 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. 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, to positive mean effects indicate beneficial conditions, 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

Table	Title
A1	Effects on Economic Self-Sufficiency and Health by Age at Randomization
A2	Effects on Employment by Age at Randomization, Administrative Data
A3	Effects on Earnings by Age at Randomization, Administrative Data
A4	Effects on Additional Economic Self-Sufficiency Outcomes
A5	Effects on Additional Health Outcomes
A6	Effects by Education Level on Economic Self-Sufficiency and Health
A7	Effects by Gender Composition of Children in the Household on Economic Self-sufficiency and Health
A8	Effects for Prime Age Minority Single-Mothers versus Others
A9	Effects on Psychological Distress, Detailed Results
A10	Effects on Voucher Use, Housing and Neighborhood Quality, and Safety
A11	Effects on Education and Training and Social Networks
A12	Effects on Health Behaviors and Health Care Access
A13	Effects on Additional Mediators: Mobility and Housing Assistance and Access to Transportation
A14	Effects on Residence in Census Tracts with Changing Characteristics between the 1990 and 2000 Censuses

Notes. Tables A4-A14 are available at: <http://www.wws.princeton.edu/~kling/mto/481a.pdf>

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 (i)	E-C ITT (ii)	S-C ITT (iii)	CM (iv)	E-C ITT (v)	S-C ITT (vi)	E-C ITT (vii)	S-C ITT (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 depression (lower bound) [SR]	.151	-.020 (.020)	-.013 (.021)	.175	-.035 (.021)	-.013 (.024)	-.014 (.029)	-.000 (.032)
Worrying [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. 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, 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.

APPENDIX TABLE A2 — EFFECTS ON EMPLOYMENT
BY AGE AT RANDOMIZATION — ADMINISTRATIVE DATA

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

Notes. ADMIN = administrative data. SR = self-report. 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. RA = random assignment. * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses. Administrative data on 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.

APPENDIX TABLE A3 — EFFECTS ON EARNINGS
BY AGE AT RANDOMIZATION — ADMINISTRATIVE DATA

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

Notes. ADMIN = administrative data. SR = self-report. 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. RA = random assignment. * = statistically significant at the 5 percent level. Standard errors, adjusted for heteroscedasticity, are in parentheses. Administrative data on 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.