The Three-City Study Incentive Experiment:

Results from the First Two Waves

Robert Moffitt
Johns Hopkins University

December, 2004

The author would like to thank Karen Jones and Richard Kulka for comments, Xiaoguo Hu for research assistance, the National Institutes of Health for financial assistance, and several members of the staff of the Research Triangle Institute for assistance on the implementation of the incentive experiment: Allen Duffer, Ralph Folsum, Bruce Jones, and Alice Turner. This paper is a product of the Three-City Study and can be found on the project Web site at http://www.jhu.edu/~welfare.

aclcrep9_v3.wpd
12/3/04

Abstract

This report presents the results of an incentive experiment in the Three-City Study which offered a random set of families a higher payment for participation. The findings indicate that response rates were increased by the higher payments, even at the initial point of successfully screening families at the door. After two waves of the survey, the response rate for the families with the higher payments was almost 7 percent greater than for those with lower payments. Differences in demographic and economic outcomes among those with different payment levels are also examined. They show few differences at the screener or first wave point, but several significant differences in demographic variables by the second wave. These differences are likely to reflect differential selection into survey completion.

The Welfare, Children, and Families Study, commonly called the Three-City Study, is a longitudinal household survey intended to track the well-being of a group of low-income families with children in the wake of welfare reform in the United States. The families were drawn from low and moderate income neighborhoods in three cities--Boston, Chicago, and San Antonio--and were first interviewed in 1999 and then again in 2000-2001. About 2,400 families were interviewed in the first wave. Most were single-mother families but a few married families with children were included as well. A large number of reports and publications have resulted from the project; see the project Web site for details, http://www.jhu.edu/~welfare.

The project included a randomized trial of the effects of incentive payments. The standard incentive payment for the interview was $30, but a random 20 percent of families were offered $70. There is a large literature on the effects of higher incentive payments on response rates to surveys which typically shows some positive response (Groves and Couper, 1998; Singer and Kulka, 2002). The incentive experiment in the Three-City Study permits another examination of this question. It is also useful because it focuses explicitly on the low-income population, for the number of incentive studies which have focused explicitly on this population is modest (Singer and Kulka, 2002). Many believe that the low income population is more responsive to incentives than the general population and some, but not all, evidence supports this (Groves and Couper, 1998, p.284; Singer and Kulka, 2002).

In addition, an incentive experiment permits, with some assumptions, an estimate of the effect of nonresponse on bias in the means of survey variables by a simple comparison of those

survey means for the members of the experimental and control groups. We use the Three-City incentive experiment data to examine this question as well.

We find that the response rate at the initial screening point was almost 3 percentage points greater for those with the higher incentive payments, and the response rate was over 6 percentage points greater at the subsequent first wave interview. Response rates were slightly higher at the second wave as well. By the end of the second wave, those with higher incentive payments had response rates almost 7 percentage points greater than the response rates of those with standard payments. The analysis of outcome variables indicates that, at the first wave, there was no significant difference between experimentals and controls in most variables--mother's education, employment status, marital status, number of children under 3, family size, earnings, household income, welfare participation and income, etc.--but there were a few variables where there were significant differences--caregiver age and Hispanic status, for example. At the second wave, there were more significant differences, some of which appeared only in unadjusted figures and others of which appeared only in adjusted differences. The variables include caregiver age, number of adults in the household, whether the mother had a high school education, marital status, cohabitation status, and health and disability.

The first section of the report provides more details on the Three-City Study and the incentive experiment. The second section of the report briefly discusses statistical issues. The third section reports the results of the analysis of the data.

The Three-City Study and the Incentive Experiment

A complete description of the sampling protocol for the Three-City Study can be found in Jones (2000) and the scientific rationale for the protocol can be found in Winston et al. (1999). A summary description of this protocol is necessary to understand the analysis of the incentive experiment data.

In the initial stage of drawing the Three-City sample, eight sets of block groups were constructed from all the blocks in the three cities in the 1990 Census (the Research Triangle Institute conducted the sample design and the survey). These eight sets were demarcated by city and by race-ethnic group.[1] The block groups within each set were ranked in descending order of the poverty rate of children in that particular city-race-ethnic group and a random sample of block groups with poverty rates above a specified cutoff were selected with selection probability proportional to size.[2] The selected block groups were divided into segments--which are areas of a size typically regarded as convenient for surveying and generally consist of 90-120 dwelling units--and a set of segments was chosen randomly from each block group. All selected segments were then counted and listed (i.e., interviewers visited the segments, counted the housing units, and recorded the addresses of all occupied dwelling units) and a random sample of dwelling units (identified by street addresses) was then selected from within each segment. There were approximately 479 segments and 40 thousand dwelling units in this sample. Each

[1] The eight were Boston Hispanics; Boston Non-Hispanic Whites; Boston Non-Hispanic Blacks; Chicago Hispanics; Chicago Non-Hispanic Whites; Chicago Non-Hispanic Blacks; San Antonio Hispanics; and San Antonio Non-Hispanic Blacks.

[2] The cutoff values of the poverty rate were 15.3 percent, 8.2 percent, and 17.1 percent in Boston, Chicago, and San Antonio, respectively (Winston et al. 1999, table A-2).

dwelling unit in the final sample had a probability of being selected equal to the product of the probabilities of being in a selected block group, of being in a selected segment, and of being in a selected dwelling unit. These probabilities will not be affected by the incentive experiment, but do imply that weights are needed to generate population-level statistics.

An attempt was then made to interview each dwelling unit and to screen that unit for eligibility for the sample. Prior to screening, an advance letter was sent notifying the occupants of the incentive payment that would be offered to them for participation in the survey. The standard incentive was $30 but a random 20 percent was offered $70. The randomization was conducted at the segment level to minimize the chance that neighbors would have different payments and would learn of this. Thus families in each segment were entirely in either the experimental or control group.

Each address was visited by an interviewer who attempted to gather the characteristics of the residents necessary to determine whether the family was eligible for the survey. The Three-City sampling protocol required that only families with one child in the age-range 0-4 or 10-14 and with income below 200 percent of the poverty level were eligible. Approximately 90 percent of dwelling units were successfully screened. This fraction could be affected by the incentive experiment because of the advance letter, although the characteristics of the families who were not screened cannot be known nor can it be known whether the incentive payment played a role in their decision.[3] The scientific goals of the study also called for oversampling of

[3] Dwelling units that were not screened, and hence were treated as nonresponse, included those with no answer at the door or where access to the building was denied, those who refused to answer the screening questions, and those for whom a language barrier prevented answering the questions (Spanish was spoken by the interviewers and hence was not a barrier). These three groups constituted roughly one-third each of the nonrespondents, with a small number in

certain families within the eligible group--those with lower income, who were likely to be on welfare, who were single mothers, and from certain race-ethnic groups. These characteristics were also gathered in the screening interview, and families with a particular set of characteristics were assigned a probability of being selected to meet these targets; once a family was determined to be eligible in the screening interview, it was invited to participate or not to participate at a rate necessary to obtain the designated fraction of families in that cell. The incentive payment was explicitly mentioned to families who were given such an offer.[4] The response rate to these offers was 87 percent. The first wave of the survey was administered immediately to those who agreed to participate, and later to those who wished to be interviewed at a later time.

Between the first and second waves of the survey (approximately 19 months), families in the experimental group were sent postcard reminders with a $5 bill included, whereas families in the control group were sent only the postcard. For the second wave itself, the incentive amounts for all families were left unchanged. The overall response rate to the second wave was 88 percent.

---

other miscellaneous categories.

[4] Screening interviewers had the option of mentioning the incentive experiment anytime during the screening interview. It is unknown how many mentioned it at the beginning of the screening interview or during it rather than waiting until the end.

Statistical and Modeling Considerations

We are interested in the effect of the incentive payment on response rates and on what we term outcome variables--that is, the characteristics of the sample reported in the first and second wave interviews. Mathematically, we denote R as a indicator variable for response (R=1 if responded, R=0 if not) and E as an indicator variable for experimental status (E=1 if in the experimental group, E=0 if in the control group) and hence our interest centers on the effect of the treatment on the probability of response, Δ_1=P(R=1|E=1) - P(R=1|E=0) where P is the probability function. We are also interested in how this treatment effect differs by individual characteristics X.

For outcome variables, we will be interested primarily in the means (i.e., first moments) of their distributions, for those who respond, rather than their variances or other distributional features. We let Y denote such an outcome variable. The distinction between X and Y is that X is a characteristic known for the entire sample, both respondents and nonrespondents, whereas Y is a characteristic known only for those who respond. Our interest will center on Δ_2 = E(Y|E=1,R=1) - E(Y|E=0,R=1), both overall and separately by X. A difference in outcomes can arise either because incentive payments affect the degree of bias in reported values given in response to a question (i.e., accuracy of response), or because incentive payments induce different types of individuals and families to participate and to respond. We will make no attempt to distinguish between these two explanations for any differences we observe. However, on the assumption that incentive payments do not much affect the accuracy of response, we may take the mean of Y in the experimental group to be a more accurate indicator of the mean of Y in the total population and the experimental-control difference to be a measure

of nonresponse bias. This inference is based on the assumption that the higher incentive brings into the sample individuals who are more like (in their values of Y) the nonrespondents than the respondents.[5] If incentive payments do not affect the accuracy of responses, and if the probability of response is independent of the value of Y (implying no bias in the first place), then there should be no experimental-control difference in the conditional mean of Y (Δ_2=0).[6]

Applying this general approach to the Three-City Study requires taking account of its multistage process. There are three events which the incentive experiment could have affected: whether the occupants of a dwelling unit completed the screener; whether those screened and offered participation in the survey agreed to participate in the first wave interview; and whether respondents at the first wave participated in the second wave interview. The combination of the first two determines overall response at the first wave. We shall denote these three stages as Stage 1A, Stage 1B, and Stage 2, respectively.

For the Stage 1A analysis of the effect of the incentive experiment on response, we can only calculate mean response rates by experimental and control segments and compare them. We cannot condition on individual characteristics X or conduct the analysis at the individual

---

[5] Groves and Couper (1998, p.13) give an example of refusal conversions where this might not be the case, for those conversions might bring in a subset of the nonrespondents who are more like the respondents, possibly increasing the bias. For incentives, this seems unlikely to occur because higher incentives are likely to bring in those who were not willing to exert the effort to respond at lower incentives and, presuming effort to be the primary reason for nonresponse in general, to bring in individuals who are more like nonrespondents. Of course, as incentives drive the response rate towards 100 percent, bias in Y eventually has to be reduced by definition.

[6] Proof: if f(Y|E,R) is the density of Y conditional on values of E and R, then Bayes' rule implies that f(Y|E,R) = [f(R|Y,E) f(Y|E) f(E)] / [f(R|E) f(E)]. If f(R|Y,E) = f(R|E) and if f(Y|E) = f(Y), then f(Y|E,R) = f(Y).

level because we have no individual or family characteristics for those not screened. However, we do collect data on area characteristics from the 1990 Census on poverty rates, female-head percents, racial composition, and related variables, and match them to the segments in our sample, thereby allowing us to use these as area X variables in our analyses. In principle they should not affect any of our estimates of experimental-control differences because they should be orthogonal to experimental-control status if the randomization was conducted effectively, but their effects on response rates are of independent interest and their inclusion may increase efficiency. For outcomes Y in the Stage 1A analysis, we use those characteristics collected in the screener: whether the family was poor or female-headed, whether it received Medicaid or Food Stamps, and the race of the household head.

For the analysis of Stage 1B, we can use the segment characteristics as well as individual characteristics collected in the screener as X variables to analyze experimental-control differences in the probability of agreeing to participate in the first wave survey unconditionally or conditional on these variables. We will examine response to the first wave survey conditional on response to the screener as well as the cumulative response starting at the segment level, which is the product of the two conditional response rates. We will analyze outcome variables, Y, for those who responded to the first wave survey, and will select traditional socioeconomic characteristics such as family structure, marital status, education, health, race, employment, earnings, and income. We will analyze these characteristics both conditional on response to the screener, in which case X variables from the screener can be used, and unconditional on that response, in which case only screener area X variables can be used.

For the analysis of Stage 2, we can use the segment characteristics, the screener

variables, and the first wave variables as X variables to analyze the probability of response at the second wave conditional on response to the first wave. Cumulative rates can be examined unconditional on the first wave and screener X variables. The outcome variables, Y, will be the same socioeconomic characteristics for the first wave but measured at the second wave.

Results

Treatment Assignment. We first determine whether the randomization was implemented effectively by determining whether it is orthogonal to variables we are able to measure. To do this we use the sample of dwelling units in the segments which were surveyed and aggregate up to the response rate in each segment--the fraction of dwelling units for whom a screening interview ("screener") was completed--and regress this rate on several observables: the size of the segment, the city-race-ethnicity stratum in which the segment was located, and the Census variables we have collected by segment.[7] The results are shown in Table 1. As the results indicate, virtually all variables are statistically insignificant and the R-squareds and F-statistics from the regressions are very small. Therefore the evidence suggests that a valid randomization took place.

Stage 1A. In this stage we examine first whether experimental-control status affected the screening response rate. Again we operate at the segment level. There are 472 segments in the analysis sample, 95 of which are experimental and 377 of whom are control (33,923

[7] The Census data are by block group, and segments are smaller than block groups, so the Census variables pertain to the larger area in which the segment is located.

individuals in total).[8] As shown in the first row of Table 2, in the experimental segments the mean response rate was 94.3 percent as compared to a 91.4 percent response rate in control segments. The 2.9 percent difference is statistically significant. Therefore it appears that there was an effect of the experimental payment at this early stage, presumably reflecting the effects of the payments promised in the advance letter.

These raw response rate comparisons were supplemented by regressions of the segment-level response rate on an experimental-control treatment dummy and the other segment-level variables used in Table 1 (segment size, strata variables, Census variables). Not surprisingly, given the results of Table 1 demonstrating the orthogonality of treatment assignment and these variables, the coefficients on the treatment dummy in these regressions remained at .029 and remained statistically significant. The fullest regression is shown in Appendix Table A-2. Most of the segment-level characteristics did not significantly affect response rates; exceptions were the percent of the population which was 65 and over and the percent of the population which was Hispanic, both of which increased response.

We also examine the effect of experimental-control status on outcomes in the screening interview. The variables available from the screener are whether the household was female-headed or poor, whether Food Stamps or Medicaid were received, the race of the household head, and whether children aged 0-4 or 10-14 were in the household. As noted previously, experimental-control differences in the means of these variables can arise either because those who receive higher payments report differently, or because the payment induces selection into

[8] A few segments were omitted because they had unequal dwelling unit weights, and a few more were omitted because they had no observations or missing data on the experimental incentive variable.

the sample which is correlated with these characteristics (we know that more individuals responded if the incentive payment was higher). Table 3 shows the experimental and control means for these variables for those who completed the screener. All differences are small and insignificant. Therefore there appears to have been no systematic bias in the response rate process at the screener level. Table 4 shows the coefficients on an experimental-control dummy variable in regressions of the screener outcome variables on that dummy and the segment-level characteristics. All coefficients save that for the presence of children in the specified age ranges remain insignficant, and the children coefficient is on the borderline of insignificance. If taken as a measure of selection into the sample, however, its positive sign implies that families with children 0-4 or 10-14 were more likely to respond if the payment was higher; this implies that the main sample, with the lower payment, has too few of such families.

Stage 1B. Among those who completed the screener and were found eligible, a certain fraction were offered participation in the first wave of the survey, as described previously. The number of families so offered were 2,954. As indicated in the second row of Table 2, 87 percent of those offered participation completed the first wave interview. Once again, more experiment families completed that interview than controls; the difference is over 6 percentage points. Thus the incentive payments continued to have a large impact on survey participation and, perhaps not surprisingly, a larger impact than at the screener stage.

The third row of Table 2 shows the cumulative response rate at the first wave taking into account nonresponse at both stages 1A and 1B. The overall response rate is about 77 percent and the experimental-control difference is a little over 4 percentage points. It is not statistically significant but, given the clear positive effects of incentive payments shown in the first two

rows, should be interpreted as a favorable effect of incentives.[9]

As before, regressions of response rates onto the experimental-control dummy but controlling for background characteristics--in this case, individual characteristics measured as of the screener and segment-level characteristics--leaves the treatment effect essentially unchanged, at 6.5 percentage point (see Appendix Table A-3). The coefficients on the other variables imply that response rates at the first wave were higher for those receiving Food Stamps or Medicaid, and higher for those residing in blocks in the San Antonio Hispanic neighborhoods than for those residing in blocks in the White neighborhoods in Boston and Chicago.

Experimental-control differences in various outcome variables measured for those who completed the first-wave interview are shown in Table 5. Simple experimental-control differences (columns labeled 'raw') show that the differences are statistically insignificant for virtually all variables. The two exceptions are the age of the caregiver, which is significant only at the 10 percent level, and an Hispanic family; taken as an effect of selection, these imply that the higher payments drew in older mothers and fewer Hispanics, suggesting that the main sample has mothers who are too young and have Hispanics in greater than population proportions. Nevertheless, the overall results suggest very few effects of these types at the first wave.

The last two columns in Table 5 show experimental-control differences after regression adjustments for background characteristics measured at the screener and at the segment level; the penultimate column uses only the dwelling unit weight whereas the final column uses weights

[9] The means for cumulative response rates shown in Table 2 do not equal the product of the first two rows because they are means taken over the products of the individual segment rates at the two stages. The mean of a set of products does not equal the product of the means. In practical terms, it appears that segments where response rates were high at the screener stage were below-average at the next stage, and vice-versa.

developed at RTI which contain a nonresponse adjustment. That adjustment accounted for differential nonresponse rates by city-race-ethnicity stratum and characteristics reported in the screener--poverty status, female headship, race-ethnicity, and receipt of Food Stamps and Medicaid (Jones, 2000). When the regression adjustments are made but without the RTI nonresponse weight adjustment, the caregiver age difference falls slightly and its standard error increases to the extent that it slightly fails at 10 percent significance test; it falls further with the RTI adjustment but so does its standard error, raising its significance once again. These results suggest that regression adjustments are not capable of eliminating the caregiver age difference arising from the incentive payment differential. The Hispanic difference falls much more dramatically in magnitude when regression adjustments are made and, while significant without the RTI weight adjustment, is only barely so. On the other hand, surprisingly, the experimental-control difference in the percent Black increases when the adjustments are made and is significant when the RTI weight adjustment is made. This implies that there must be race differences in response within the cells of city-race-ethnicity strata and among groups with the same poverty rate, welfare receipt, and family headship type (i.e., the variables controlled for in the regression).[10] Nevertheless, these race differences are still isolated cases when considered against the results in the rest of the table for the other variables. In addition, most research applications to which the Three-City data have been put do not have race as a dependent variable but rather use it as a control variable, in which case these response rate differences will not affect

[10] It should be noted that response rate differences by time-invariant characteristics such as race can be completely eliminated by weighting because race is known with certainty from the screener, assuming no inaccuracies in response. The result for the RTI-adjusted difference implies that that adjustment is not being made within the cells controlled for in the regression.

the estimates of any such research models.

Stage 2. About 2,437 families completed the first wave interview were eligible for completing the second wave interview. As shown in Table 2, the conditional second wave response was 88 percent. However, it was only slightly higher for experimental families than for control families (2.2 percentage point difference), suggesting that the effect of the incentive payments may be falling over time. It is possible that those still in the survey by the second wave were relatively easy to find, or committed to participation, resulting in less of an effect of monetary incentives across the remaining sample families. Regression adjustments to this second-wave response rate made little difference to the experimental-control difference, which remains low and statistically insignificant (see Appendix Table A-4).

The cumulative response rate shown in Table 2 is 66 percent overall and almost 7 percentage points higher for experimental families than for controls. While the standard error of the treatment effect has risen and renders it below significance at the conventional 10 percent level, it is clear that the magnitude of the difference remains sizable after two waves.

The results for second-wave outcome variables are shown in Table 6. A greater number of significant differences now appear. Taken as an indication of selection, the results for raw differences suggest that the higher experimental payments brought in more unmarried and cohabiting caregivers, more women in poor health or disabled, and, once again, more non-Hispanic households. While these demographic variables are now affected by incentive payments, all economic variables remain unaffected. As for the regression adjustments, both those with and without the RTI nonresponse adjustment have a mixed pattern of results. For the some of the significant raw differences, the adjustment does not eliminate the difference

(cohabiting, poor health); for others, the significance is eliminated by the RTI adjustment but not otherwise (married, disabled); while for yet others, the differences become significant when the regression adjustments are made (caregiver age, number of adults, high school education).[11] The RTI nonresponse adjustment should not necessarily be expected to eliminate significance of differences because the demographic variables under discussion were not used as factors in the nonresponse adjustment.

The sensitivity of the results for some of these characteristics, particularly those for marriage and cohabitation, educational attainment, and health, should be a source of concern because those variables are used not infrequently as dependent variables in research uses of the Three-City data. Taken as a reflection of selection, the signs on these variables imply that the main sample may have too few unmarried, non-cohabiting caregivers, too few women with low or high educational attainment (most likely the former because it is the majority of the sample), and too few women who are disabled or in poor health. Together, these three types of characteristics suggest that second-wave nonresponse was disproportionate among the more disadvantaged.[12]

The implications of these findings for research use of the Three-City data cannot be easily drawn because most research models are considerably more complex than the simple least-squares regressions here, and generally specify a different set of independent variables. In

[11] The race results are the same as for the first wave.

[12] The regressions include only the same screener-level and segment-level variables used in the first wave outcome regressions. Additional regressions were estimated which included the first-wave values of the dependent variables in each of these regressions. The experimental-control differences were generally unchanged as a result.

addition, of course, many studies use outcome variables not examined here. However, the availability of the incentive experiment makes it possible for any research study to estimate the effect of nonresponse on its particular model estimates by making use of an experimental-control incentive dummy and by testing whether the coefficients in a model are different for the two groups. Such a set of tests would be a useful check for nonresponse bias.

Summary

This report presents findings from an incentive experiment conducted as part of the Three-City Study which offered a substantially higher monetary reward for participation to a random set of families. The results indicate that the payments affect response rates favorably at the initial screening point, at the first wave, and at the second wave. By the end of the second wave, response rates among those with the higher payments were almost 7 percentage points greater than those with the standard payments.

The report also presents differences in outcome variables at the different stages of the survey. While few differences are found at screener or first-wave stage, several demographic differences are found at the second wave. The size and significance of these differences is sensitive to the whether regression adjustments are made and whether nonresponse weights are utilized. It is suggested that future uses of the Three-City data test the sensitivity of their estimates of raw differences and model coefficients to nonresponse bias by using the incentive experiment variation.

References

Groves, R. and M. Couper. 1998. Nonresponse in Household Interview Surveys. New York: Wiley.

Jones, B.L. 2000. "Welfare, Children, and Families: Sample Selection and Weight Development." Research Triangle, North Carolina: Research Triangle Institute.

Singer, E. and R. Kulka. 2002. "Paying Respondents for Survey Participation," In Studies of Welfare Populations: Data Collection and Research Issues, eds. Michele Ver Ploeg, Robert A. Moffitt, and Constance F. Citro, pp. 105-128. Washington: National Academy Press.

Winston, P., R. Angel, L. Burton, P. L..Chase-Lansdale, A. Cherlin, R. Moffitt, and W. J. Wilson 1999. Welfare, Children, and Families: A Three-City Study, Overview and Design. Johns Hopkins University. Available at www.jhu.edu/~welfare.

Table 1

Segment-Level Treatment Assignment Regressions

Variable	(1)	(2)	(3)
Segment size	-.000 (.001)	-.000 (.001)	-.000 (.001)
City-race-ethnicity:			
Boston White	--	.027 (.102)	.230 (.201)
Chicago White	--	.017 (.086)	.196 (.161)
Boston Black	--	-.012 (.124)	.183 (.227)
Chicago Black	--	-.024 (.099)	.127 (.230)
San Antonio Black	--	.024 (.103)	.014 (.147)
Boston Hispanic	--	.009 (.098)	.229 (.177)
Chicago Hispanic	--	.012 (.100)	.141 (.161)
Census Area Variables			
Percent Black	--	--	.052 (.181)
Percent Hispanic	--	--	.070 (.188)
Percent female-headed families	--	--	-.042 (.376)
Percent of children in female-headed households	--	--	-.389 (.262)

Table 1 (continued)

Variable	(1)	(2)	(3)
Percent women married	--	--	-.520 (.582)
Percent persons 65+	--	--	1.013 (.711)
Percent persons with fewer than 12 years of education	--	--	-.288 (.509)
Median household income (divided by 1000)	--	--	-.010 (.006)
Percent households without earnings	--	--	-.810* (.478)
Percent households with public assistance income	--	--	.078 (.325)
Percent families who are poor	--	--	-.601 (.514)
Percent households who are poor	--	--	.956 (.716)
Constant	.193** (.069)	.213** (.079)	.613 (.400)
R-squared	.000	.002	.052
F-statistic	.01	.07	.95

Notes:
Data are at the segment level (N=472) and weighted by segment size and dwelling unit weight. Dependent variable = fraction of observations in segment in the experimental group
Segments with zero observations, with unequal dwelling unit weights, and with missing treatment variables are omitted (7 segments). Means of the variables are shown in Appendix Table A-1. Omitted city-race-ethnicity group is San Antonio Hispanics. Standard errors in parentheses; **: significant at the 5 percent level; *: significant at the 10 percent level.

Table 2

Response Rates by Stage and Overall

	Total	Experimental	Control	Difference
Stage 1A (screener response rate)	.920	.943	.914	.029** (.011)
Stage 1B (first wave response rate among screener respondents)	.870	.919	.858	.061** (.022)
Cumulative First Wave Response Rate	.768	.803	.760	.043 (.038)
Stage 2 (second wave response rate among first wave respondents)	.880	.897	.875	.022 (.033)
Cumulative Second Wave Response Rate	.661	.717	.648	.069 (.044)

Notes:
Stage 1A and the two cumulative figures are computed at the segment-observation level and are weighted by segment size and dwelling unit weight. Conditional Stage 1B and Stage 2 figures are computed at the individual level and are weighted by the dwelling unit weight only. Note that the cumulative figures do not equal the product of the conditional figures partly because of differences in weights but mostly because the cumulative rates are computed at the individual segment level and then averaged, rather than being computed simply as the product of the average conditional rates.
Figures in parenthesis are standard errors
**: significant at the 5 percent level

Table 3

Experimental and Control Means of Screener Variables

	Total	Experimental Group	Control Group	Experimental-Control Difference
Female-headed household	.593	.606	.590	.016 (.031)
Poor	.673	.659	.677	-.108 (.193)
Received Food Stamps or Medicaid	.522	.491	.529	-.037 (.035)
Race:				
Black	.513	.534	.507	.027 (.034)
Hispanic	.450	.415	.459	-.043 (.031)
White	.038	.050	.034	.016 (.012)
Children 0-4 or 10-14 present	.307	.323	.303	.019 (.031)

Notes:
N = 30,865
Sample consists of all individuals who completed a screening interview
Dwelling unit weights used
Standard errors in parentheses

Table 4

Treatment Coefficients in Screener Variable Regressions

Outcome Variable	Treatment Coefficient
Female-headed household	.017 (.028)
Poor	-.020 (.031)
Received Food Stamps or Medicaid	-.027 (.033)
Black	.004 (.011)
Hispanic	-.018 (.007)
White	.015 (.009)
Children 0-4 or 10-14 present	.027* (.016)

Notes:
N = 30,865
Sample consists of all individuals who completed a screening interview
Other variables in the regression: all city-ethnicity screening variables and Census area variables shown in Table 1
Dwelling unit weights used
Standard errors in parentheses. *: significant at the 10 percent level.

Table 5

Treatment Impacts on First Wave Outcome Variables

Outcome Variable	Control Group Mean	Experimental-Control Difference		
		Raw	Adjusted	Adjusted, RTI Weights
Demographic Characteristics				
Mother age	32.6	2.69*	2.09	1.92**
		(1.56)	(1.48)	(.98)
No. adults	1.73	-.106	-.082	-.076
		(.093)	(.095)	(.068)
No. children	2.86	.111	.161	.017
		(.183)	(.179)	(.152)
Presence of children under 3	.514	.083	.093	-.004
		(.056)	(.058)	(.048)
Household size	4.59	.005	.080	-.053
		(.209)	(.202)	(.159)
Education:				
Less than high school	.428	-.052	-.035	-.019
		(.057)	(.056)	(.042)
High school	.351	.074	.060	-.009
		(.059)	(.056)	(.043)
More than high school	.221	-.022	-.025	.027
		(.045)	(.044)	(.053)
Married	.216	-.041	.000	-.045
		(.043)	(.039)	(.054)
Cohabiting	.051	-.010	-.001	-.012
		(.019)	(.018)	(.018)

Table 5 (continued)

Outcome Variable	Control Group Mean	Experimental-Control Difference		
		Raw	Adjusted	Adjusted, RTI Weights
Disabled	.188	.058	.058	-.001
		(.052)	(.053)	(.029)
Poor health	.265	.065	.073	.051
		(.056)	(.055)	(.053)
Black	.717	.003	-.049	-.046**
		(.043)	(.033)	(.026)
Hispanic	.258	-.074**	-.016*	-.019
		(.028)	(.009)	(.025)
White	.019	.028	.026	.027
		(.021)	(.021)	(.018)
Economic Variables				
Employed	.476	-.038	-.050	-.002
		(.059)	(.058)	(.051)
Mother earnings	406.2	-46.6	-69.4	22.3
		(60.0)	(56.7)	(56.7)
Other Household earnings	175.3	-4.37	2.68	-27.6
		(51.1)	(46.5)	(69.9)
On welfare	.486	.029	.047	-.007
		(.059)	(.051)	(.029)
Mother welfare income	157.4	2.02	5.18	-3.94
		(21.4)	(19.12)	(11.4)
Total household income	1213.9	-62.0	-87.4	-31.3
		(74.5)	(73.4)	(66.3)

Notes: N=2,366. "Adjusted" results are from regressions with the additional independent variables listed in Table A-3; black and hispanic variables are omitted from the regressions for the three race variables. All incomes are monthly. Standard errors in parentheses. **: significant at the 5 percent level. *: significant at the 10 percent level.

Table 6

Treatment Impacts on Second Wave Outcome Variables

Outcome Variable	Control Group Mean	Experimental-Control Difference		
		Raw	Adjusted	Adjusted, RTI Weights
Demographic Characteristics				
Mother age	34.3	2.51	1.79	2.94**
		(1.66)	(1.58)	(1.34)
No. adults	1.86	-.327	-.316**	-.183**
		(.088)	(.089)	(.076)
No. children	2.86	.038	.114	.089
		(.188)	(.186)	(.192)
Presence of children under 3	.442	.008	.037	-.024
		(.061)	(.061)	(.059)
Household size	4.72	-.290	-.202	-.095
		(.212)	(.209)	(.214)
Education:				
Less than high school	.432	.054	.080	.052
		(.062)	(.059)	(.054)
High school	.454	-.094	-.112**	-.052
		(.060)	(.056)	(.049)
More than high school	.114	.040	.033	-.001
		(.046)	(.045)	(.030)
Married	.274	-.138**	-.078**	-.034
		(.036)	(.034)	(.034)
Cohabiting	.090	-.042**	-.039*	-.044**
		(.019)	(.020)	(.020)

Table 6 (continued)

Outcome Variable	Control Group Mean	Experimental-Control Difference		
		Raw	Adjusted	Adjusted, RTI Weights
Disabled	.193	.129**	.123**	.052
		(.058)	(.057)	(.040)
Poor health	.262	.144**	.143**	.131**
		(.059)	(.060)	(.050)
Black	.715	.020	-.032	-.075**
		(.043)	(.028)	(.037)
Hispanic	.259	-.073**	-.015	.003
		(.030)	(.009)	(.026)
White	.022	.005	.003	.022
		(.007)	(.007)	(.023)
Economic Variables				
Employed	.571	.029	.021	.015
		(.059)	(.060)	(.057)
Mother earnings	646.8	-33.4	-68.4	-28.1
		(91.4)	(86.3)	(68.6)
Other Household earnings	379.4	-103.0	-83.1	55.9
		(68.1)	(62.9)	(87.6)
On welfare	.282	.057	.076	.032
		(.059)	(.057)	(.039)
Mother welfare income	78.4	12.7	13.7	12.6
		(17.4)	(17.4)	(12.6)
Total household income	1576.4	65.7	12.4	208.8
		(166.5)	(140.6)	(188.6)

Notes: N=2,103. "Adjusted" results are from regressions with the additional independent variables listed in Table A-3; black and hispanic variables are omitted from the regressions for the three race variables. All incomes are monthly. Standard errors in parentheses.
**: significant at the 10 percent level. **: significant at the 5 percent level.

Appendix Table A-1

Means of the Variables in Table 1

Experimental-control indicator	.189
Segment size	71.1
City-race-ethnicity:	
Boston White	.057
Chicago White	.174
Boston Black	.036
Chicago Black	.351
San Antonio Black	.046
Boston Hispanic	.049
Chicago Hispanic	.101
Census Area Variables	
Percent Black	.456
Percent Hispanic	.336
Percent female-headed families	.169
Percent of children in female-headed households	.400
Percent women married	.283
Percent persons 65+	.083
Percent persons with fewer than 12 years of education	.269

Appendix Table A-1 (continued)

Percent households without earnings	.327
Percent households with public assistance income	.215
Percent families who are poor	.431
Percent households who are poor	.421

Appendix Table A-2

Segment-Level Screener Response Rate Regression

Independent Variable	
Experimental dummy	.028**
	(.010)
Segment size	-.000
	(.000)
City-race-ethnicity:	
Boston White	.013
	(.055)
Chicago White	-.006
	(.038)
Boston Black	.002
	(.055)
Chicago Black	.016
	(.049)
San Antonio Black	.040
	(.027)
Boston Hispanic	-.014
	(.045)
Chicago Hispanic	.028
	(.035)
Census Area Variables	
Percent Black	.010
	(.035)
Percent Hispanic	.093**
	(.043)
Percent female-headed families	-.041
	(.065)

Appendix Table A-2 (continued)

Independent Variable	
Percent of children in female-headed households	.061 (.046)
Percent women married	.021 (.132)
Percent persons 65+	.411** (.174)
Percent persons with fewer than 12 years of education	-.126 (.118)
Median household income (divided by 1000)	.000 (.001)
Percent households without earnings	-.070 (.071)
Percent households with public assistance income	-.042 (.047)
Percent families who are poor	.222 (.145)
Percent households who are poor	-.100 (.142)
Constant	.828** (.082)
R-squared	.191
F-statistic	7.07

Notes:
Data are grouped at the segment level (N=472). Dependent variable = response rate in a segment (mean = .920). Segments with zero observations, with unequal dwelling unit weights, and with missing treatment variables are omitted (7 segments). Observations weighted by segment size and dwelling unit weight. Omitted city-race-ethnicity group is San Antonio Hispanics. Standard errors in parentheses. **: significant at the 5 percent level.

Appendix Table A-3

Conditional First Wave Response Rate Regression

Independent Variable	Coefficient
Experimental-control dummy	.065**
	(.023)
Female-headed household	-.021
	(.031)
Poor	.009
	(.034)
Received Food Stamps or Medicaid	.119**
	(.037)
Black	.030
	(.030)
Hispanic	-.042
	(.031)
Segment size	-.000
	(.000)
City-race-ethnicity:	
Boston White	-.062*
	(.034)
Chicago White	-.108**
	(.038)
Boston Black	-.049
	(.044)

Appendix Table A-3 (continued)

Independent Variable	Coefficient
Chicago Black	-.027 (.047)
San Antonio Black	.006 (.034)
Boston Hispanic	-.016 (.034)
Chicago Hispanic	-.045 (.033)
Constant	.809** (.054)
R-squared	.048
F-statistic	8.20

Notes:

N = 2,954
Sample consists of all individuals who completed a screening interview
Dependent Variable = 1 if completed first wave interview, =0 if not
Dwelling unit weights used
Omitted city-race-ethnicity group is San Antonio Hispanics.
Standard errors in parentheses. **: significant at the 5 percent level. *: significant at the 10 percent level.

Appendix Table A-4

Conditional Second Wave Response Rate Regression

Independent Variable	Coefficient
Experimental-control dummy	.031 (.032)
Female-headed household	.010 (.043)
Poor	-.004 (.032)
Received Food Stamps or Medicaid	-.097** (.021)
Black	.147 (.155)
Hispanic	.123 (.129)
Segment size	.000 (.000)
City-race-ethnicity:	
Boston White	-.007 (.045)
Chicago White	-.103** (.040)
Boston Black	-.121 (.044)

Appendix Table A-4 (continued)

Independent Variable	Coefficient
Chicago Black	-.029 (.050)
San Antonio Black	-.026 (.033)
Boston Hispanic	-.033 (.030)
Chicago Hispanic	-.239 (.037)
Constant	.832** (.131)
R-squared	.044
F-statistic	11.21

Notes:

N = 2,437
Sample consists of all individuals who completed the first wave interview
Dependent Variable = 1 if completed second wave interview, =0 if not
Dwelling unit weights used
Omitted city-race-ethnicity group is San Antonio Hispanics.
Standard errors in parentheses. **: significant at the 5 percent level.