

# Achieving MTO's High Effective Response Rates: Strategies and Tradeoffs

Nancy Gebler  
University of Michigan

Lisa A. Gennetian  
The Brookings Institution

Margaret L. Hudson  
University of Michigan

Barbara Ward  
University of Michigan

Matthew Sciandra  
National Bureau of Economic Research

*The contents of this chapter are the views of the authors and do not necessarily reflect the views or policies of the U.S. Department of Housing and Urban Development, the Congressional Budget Office, the U.S. government, or any state or local agency that provided data.*

---

## Abstract

*The Institute for Social Research (ISR) at the University of Michigan successfully led an intensive, long-term, in-person survey for the Moving to Opportunity (MTO) for Fair Housing demonstration final impacts evaluation (Sanbonmatsu et al., 2011), achieving final effective response rates (ERRs) of 89.6 percent among MTO adults and 88.7 percent among youth, well above what response rates of surveys with comparable low-income populations have accomplished. A variety of survey field strategies ISR employed—careful staff selection, strategic use of financial incentives, and close collaboration between ISR and the National Bureau of Economic Research—all contributed to these high ERRs. The high costs associated with achieving high ERRs for in-person surveys like that employed in MTO raises questions about added value. Costs per survey interview nearly quadrupled during the last 4 fielding months. This extra investment increased the MTO adult survey ERR by only about 3.2 percentage points. A reanalysis of intention-to-treat estimates on selected outcomes suggests the merits of such an investment. If survey fielding had stopped at an 81-percent ERR for adults, we would have falsely concluded that MTO had no effect on two of four key health outcomes, that MTO had no effect on female youth mental health, and that MTO increased female youth idleness.*

## Introduction

As early as the 17th century, scientists observed that individuals who live in economically disadvantaged neighborhoods fare worse on a range of outcomes—from physical and mental health to employment and earnings, schooling, crime, and consumer bankruptcy filings—than individuals who live in economically well-off neighborhoods (Macintyre and Ellaway, 2003). Untangling whether neighborhoods per se, or the variety of characteristics of the individuals residing in particular neighborhoods, drive this observed association has been difficult. The question of whether neighborhoods matter is further complicated by the fact that we cannot always observe or measure the reasons why individuals decide to live in particular types of neighborhoods, and these very same reasons might be highly related to their outcomes. The Moving to Opportunity (MTO) for Fair Housing demonstration is a study uniquely positioned to contribute to our understanding of whether neighborhoods have causal effects on individuals' well-being. The MTO experiment produced changes in housing mobility and subsequent experiences in low-poverty neighborhoods that can help isolate the effects of neighborhood circumstances on outcomes from a host of other individual, household, or local community characteristics.

To maximize MTO's contribution to science and policy, the long-term survey for the final impacts evaluation<sup>1</sup> (Sanbonmatsu et al., 2011) had an ambitious data collection strategy that included a broad set of outcomes measured from administrative records sources and an intensive in-person survey, occurring up to 15 years after study entry, led by the Institute for Social Research (ISR) at the University of Michigan. The National Bureau of Economic Research (NBER) research team set very high response rate goals to ensure that survey data collection adequately represented the eligible MTO sample and captured a breadth of outcomes across the domains of housing, neighborhood safety, physical and mental health, employment, education, financial security, and youth risky behavior. ISR successfully reached a final effective response rate (ERR)<sup>2</sup> of 89.6 percent among MTO adults and 88.7 percent among MTO youth (ages 10 to 20 as of December 2007) in the long-term survey for the final impacts evaluation. These ERRs are much greater than what some studies of low-income populations have accomplished (Weiss and Bailar, 2002) and on par with several long-established and well-resourced survey initiatives such as the Panel Study of Income Dynamics<sup>3</sup> (American Association of Public Opinion Research, 2012; Gouskova, 2008; Groves et al., 2004).

Reaching such high response rates not only required substantial time investment from the NBER research team (to fundraise and design the survey) and financial commitment from the U.S. Department of Housing and Urban Development (HUD) and various other funders, but it also required creativity and flexibility among ISR staff to navigate and strategize in real time while

---

<sup>1</sup> Research on MTO originally launched separately for each site, with a series of academic research investigators leading each site. The followup survey for the MTO interim impacts evaluation (Orr et al., 2003) that Abt Associates Inc. conducted was the first effort to administer a comparable data collection for MTO families overall. HUD also funded Abt Associates to canvass MTO families through 2007 to maintain an updated contact list.

<sup>2</sup> Ludwig (2012) describes the calculation of the effective response rate, which reflects the weighted proportion of interviews completed for the eligible adult and youth samples.

<sup>3</sup> The Panel Study of Income Dynamics (PSID) obtains response rates of 93 to 94 percent, and recent studies of PSID youth have response rates of between 87 and 91 percent.

interviewers worked in the field. Following and finding thousands of economically disadvantaged families who lived or currently live near resource-poor, potentially unsafe neighborhoods is a complex task. This complexity was compounded by the amount of time that had passed since the last in-person or phone contact with MTO study members and changes in MTO households—many of the youngest cohort at MTO study entry have since split off to create their own households. The previous in-person interview with an MTO study household member had been a minimum of 5 years before the start of the long-term survey data collection for the final impacts evaluation, and some of the sample (37 percent, or 3,830) had not been interviewed at the followup survey for the interim impacts evaluation (Orr et al., 2003), meaning the most recent contact may have occurred more than 10 years before the start of this data collection. In addition to facing the pure locational challenges of finding the eligible MTO survey sample in light of the high overall ERR aims, MTO researchers wanted to maintain balance in the temporal flow of completed interviews by site and by treatment status. Maintaining sample balance in this way required particular monitoring, nimbleness, and flexibility among ISR's data collection staff to target eligible survey sample members strategically, on a week-by-week basis, for extra attention from interviewers.

Some challenges that the ISR data collection staff faced for the MTO long-term survey effort are common to survey data collection efforts in general, but many were relatively unique or new to the experiences of ISR. The first sections of this article describe the various data collection design strategies ISR employed to maximize the probability of achieving the high ERR and strategies ISR implemented to address unanticipated challenges that protected, as much as possible, the quality of data and research design after survey data collection was in the field. These sections address factors that contributed to (and worked against) achieving a high response rate for both adults and youth.

Overall, the MTO in-person long-term survey for the final impacts evaluation was a reliable, efficient, and essential resource for capturing multiple aspects of life circumstances and individual outcomes that otherwise would have been difficult to capture at scale compared with lower cost alternatives. As one very poignant example, researchers would not have discovered MTO's surprising effects on mental health outcomes if not for a survey instrument with diagnostic questionnaires used to measure mental health disorders. The intensive efforts required to achieve the very high MTO response rates do raise questions, however, about the relative worth of extra resources necessary to complete interviews among those last, difficult-to-find respondents. If NBER researchers and the MTO study funders had spent fewer resources and stopped data collection at a lower response rate, what would the estimated effects on survey-based outcomes have looked like? In an attempt to evaluate the ex post scientific value of expending additional resources on increasing the survey response rate in a study such as MTO, the last section of this article describes the cost of MTO long-term survey data collection and a few back-of-the-envelope calculations of MTO's effects under varying response rate assumptions.

## **Background**

The MTO demonstration began in the mid-1990s at five sites (Baltimore, Boston, Chicago, Los Angeles, and New York City). Low-income families with children living in public housing in highly disadvantaged areas who volunteered for the MTO program were randomly assigned to one of

three groups: an experimental group that was offered housing vouchers that had to be used in a low-poverty area along with mobility counseling from nonprofit agencies, a Section 8 group that was offered a traditional housing voucher with no locational restrictions, and a control group that was not offered a housing voucher but remained eligible for any public assistance to which they were otherwise entitled.

MTO long-term survey fielding launched in June 2008 and continued through April 2010, with a carefully staged release of sample by each of the initial five MTO sites across three waves,<sup>4</sup> with second-stage subsampling of the hardest-to-locate cases triggered at a predetermined initial response rate threshold of 75 percent. The ISR data collection staff conducted interviews in person, using a laptop computer and averaging 108.3 minutes for adult interviews and 116.7 minutes for youth interviews. The staff interviewed one adult and up to three youth ages 10 to 20 in each MTO family. In families with more than three eligible youth, ISR randomly selected three for inclusion in the sample. In addition to employing computer-assisted interviews, the data collection protocol included taking physical measurements, collecting dried blood spot samples from adults, and facilitating achievement assessments for youth.

Conducting an extensive and complex data collection effort with a highly disadvantaged and mobile population posed many challenges. First locating the family; then convincing respondents to participate; and finally completing a survey, physical measurements, and achievement assessments on multiple individuals in a single household combined to make this data collection operation extremely difficult. Some MTO families included foster children or youth who left home several years before the interview and had not kept in contact with other family members. Address information was often outdated or incorrect, and many families were living under the radar, without credit cards, mortgages, driver's licenses, or other identification that could help locate respondents. As a result, interviewers often conducted tracking with a labor-intensive, door-to-door search, checking address information and asking neighbors if they had information about where the respondent or family may have moved.

In addition to tracking challenges, the ISR data collection staff faced considerable challenges working in and around economically disadvantaged neighborhoods. The interviewing protocol required interviewers to carry a laptop computer and a large bag of supplies and equipment.<sup>5</sup> Many areas had no public parking, requiring interviewers to carry the equipment long distances in inclement weather and up multiple flights of stairs in highrise buildings. Interviewers often had to work in unsafe neighborhoods and conducted many interviews in suboptimal locations, including small and crowded living rooms, sometimes with no heat or electricity. Family members and friends coming and going, loud televisions and radios, and other distractions often made it difficult to maintain respondents' focus and confidentiality. Finally, ISR experienced staffing shortages in some areas because of the inability to recruit and retain qualified interviewers willing and able to

---

<sup>4</sup> The first release of sample was in June 2008, the second in September 2008, and the third in February 2009, upon securing enough funding to survey a random two-thirds of Section 8 group adults.

<sup>5</sup> The combined weight of the laptop computer, equipment, and supplies needed for completing an interview was approximately 30 pounds.

work successfully under the challenging conditions MTO required. As a result, production progress across the five MTO cities was at times uneven and required adjustments in field protocols to take into account differences in completion rates by site.

## **MTO Survey Data Collection: Strategies for Optimizing Response Rates**

ISR used a multifaceted set of approaches to address the previously listed challenges. This section discusses the ISR field structure and tracking efforts and the tools and strategies it used to maximize respondent participation.

### **Field Team Structure**

In any study, building an effective field team is an essential element of successful data collection. ISR developed its field staffing model for MTO to take advantage of the clustered sample and to address the challenges of working with a sample that was highly mobile and lived in disadvantaged areas. The ISR data collection staff was made up of seven teams: a team of approximately 8 to 20 field interviewers in each of the five MTO cities, a team of Internet trackers supporting the field interviewers, and a travel team of field interviewers. The tracking team was composed of individuals who had Internet access to public records and were skilled at conducting Internet searches, networking, and piecing together information from many different sources. The travel team comprised experienced field interviewers with a demonstrated ability to work effectively and efficiently in the field and who lived in other parts of the country and were available to travel to interview respondents who had moved away from the main MTO cities. The travel team also supplemented data collection in cities that were understaffed.

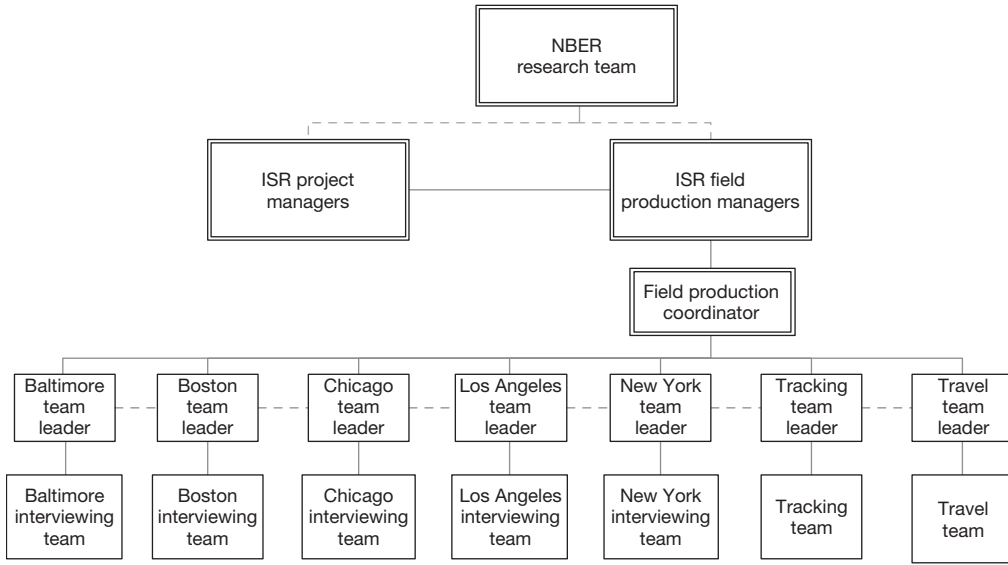
A team leader, who provided coaching and monitored quality and production and efficiency statistics, supervised each team. A production manager in the central office worked with a production coordinator in the field to supervise the team leaders and provide guidance and direction for the field activities, lead training activities, and act as the main conduit for the flow of information between the central office and the dispersed MTO field team. Exhibit 1 depicts the field staffing structure.

### **Interviewer Hiring and Training**

ISR recruited and hired all data collection staff—field interviewers, travelers, tracking experts, and supervisory staff. When recruiting staff for MTO, making sure the field staff members could work safely in low-income neighborhoods was important. The production manager and human resources specialist worked together with local interviewers already on staff at ISR to identify local advertising resources, such as neighborhood newspapers, community centers, libraries, and churches. The job posting identified some areas of the city in which MTO interviewers would work, clearly outlined the challenges interviewers would face working on this project, and emphasized the importance of the study. This emphasis on hiring local interviewers who were familiar with MTO neighborhoods and comfortable working in disadvantaged areas proved challenging (recruitment goals were not met in all areas), but it also yielded an exceptionally dedicated data

**Exhibit 1**

**ISR Field Team Structure**



*ISR = Institute for Social Research (University of Michigan). NBER = National Bureau of Economic Research.*

collection staff who remained with MTO until the end of the project. Approximately 75 percent of the ISR interviewers were newly hired (some had previous social science interviewing experience with other organizations but were new to ISR). All of the field supervisory staff (team leaders, production coordinator, and production manager) were experienced staff members who had worked on a variety of other social science data collection studies at ISR.

ISR conducted two separate training sessions for field interviewers in June and September 2008. Although having two training sessions was not the initial plan, it turned out to be beneficial in many ways. The smaller groups enabled team leaders to work more closely with each interviewer, and the slower start to production in June provided a good shakedown period, during which ISR identified and resolved issues and made adjustments ahead of the later training. Having two recruitment and training periods also enabled ISR to capitalize on the enthusiasm for the project of the first group of interviewers to help recruit additional field interviewers. Finally, the ISR data collection staff trained during the first session shared valuable tips and tricks with the newer team members, helping to mentor new interviewers and bring them up to speed more quickly after training.

Retention of the field staff was higher for MTO than for similar ISR studies. Of the 91 interviewers and trackers trained and certified to work on MTO, 85 percent were still on staff 6 months after the start of data collection, and 74 percent remained on staff through the first year of data collection. Starting in June 2009, ISR began intentionally consolidating the interview sample and releasing some interviewers to work on other projects as the second-stage subsampling began and the available number of respondents was reduced. Even with this reduction in the second half of 2009, 50 percent of the field staff stayed with MTO through the end of 2009.

## **Communication**

The central office project managers and production manager, in consultation with the NBER research team, set data collection goals and priorities for the data collection staff in the field. Each week, the ISR project managers met with the NBER research team to discuss progress, issues, requests from the field, strategies for improvement, and priorities. The priorities and action steps that emerged from this meeting provided direction for the following week's field activities. The close collaboration among the ISR project managers, the NBER research team, and the field supervisors on the data collection staff resulted in a common set of goals and priorities, which were relayed to the field teams through standardized weekly meeting agendas and reinforced in a weekly newsletter, continuous training sessions, and coaching for groups and individuals.

In addition, field supervisors on the data collection staff met weekly with the ISR project managers to review progress, learn of project updates and priorities from the NBER research team, and together develop a set of standardized agenda items for the weekly team meetings. Each of the seven data collection teams met weekly and reviewed the project agenda items and discussed team-specific issues and items. The ISR project managers produced a weekly newsletter that reinforced the training and discussions in the weekly meetings. The newsletter started as a tool to help connect the ISR project managers to the field, focusing on project updates and reminders from the ISR central office. It was adapted over time to include notes of thanks and congratulations, progress updates, training reminders, tips for effective interviewing, and more. Communication flowed from the data collection staff in the field to the ISR central office and NBER through weekly written field reports and discussions in weekly management meetings. This communication structure often led to adjustments or changes in procedures and materials based on recommendations from team leaders and interviewers. Communication was also very important from a safety perspective. ISR provided interviewers and team leaders with cell phones and asked them to let their team leader know when they were interviewing alone in dangerous neighborhoods, in unfamiliar areas, or after dark.

## **Team Interviewing**

The use of team interviewing was another feature of the MTO data collection. Because the MTO protocol included interviewing multiple individuals in a family, ISR encouraged team interviewing, whereby two (or occasionally three) interviews were conducted in a household at the same time. This procedure had several benefits: it reduced the time burden on families when multiple interviews were conducted; improved interviewer efficiency; and reduced safety concerns for interviewers traveling into dangerous areas, especially for evening interviews. It also helped ensure confidentiality by keeping the parent occupied with her<sup>6</sup> own interview while the youth was being interviewed. Finally, it enabled experienced interviewers to mentor others, helped create strong local teams that supported each other, and fostered the sharing of strategies for effective data collection. In addition to participating in team interviewing, teams often worked together on tracking blitzes, in which all members of a local team would focus on tracking for a weekend, working closely with the tracking team to try to locate as many individuals as possible and set up interview appointments for the following week. Tracking was a difficult and often frustrating part of the process, and having the

---

<sup>6</sup> Most adult respondents were female (although some male adults were in the sample).

entire team track respondents together and share success stories proved to be very motivating and an effective way to complete interviews. Likewise, interviewers who were effective at tracking could pair with those who were not as strong to help teach valuable skills and techniques.

## **Tracking and Locating Respondents**

ISR integrated tracking into its data collection operations from the start of the project and developed sample management systems to facilitate both the interviewing process and tracking lost respondents. To help limit the number of contacts needed to locate each family member, take advantage of family connections for making multiple interview appointments, and locate respondents who had moved, ISR loaded all addresses (including those at study entry, those at the time of the interim MTO study, and updates obtained through the National Change of Address system and from the HUD database) into the sample management system for each family. ISR assigned all members of a family to one interviewer, making that single interviewer responsible for locating the family. The goals were to reduce burden on the family by using one point of contact and to gather information for multiple family members in a single contact. Using this procedure served to contain costs by working first with a family-locating instrument, which enabled the interviewer to contact any family member (or informant who knew where a family or respondent may be located) and update contact information for all selected members of the family. The family-locating instrument also gave the interviewer a view of the entire work scope for that family: the number of interviews to be completed at a single address and the ages of the respondents (which determined, for example, which youth achievements assessments would be administered). After the family instrument was complete, individual sample lines were released to the interviewer with contact information updated from the family instrument, enabling each individual survey respondent to be contacted, tracked if necessary, and interviewed. Often, interviewers were able to set up the appointments for interview sessions (and arrange for team interviewing) when working with the family instrument.

In addition to taking advantage of family connections to help locate respondents who had moved since the last interview, the data collection staff included a tracking team that was integrated into the five local interviewing teams (one at each MTO site) and the travel team. Although all trackers could and did work sample lines from any location, having one primary tracker working with each local interviewing team enabled the tracker to establish strong working relationships with the field interviewers and to become very familiar with the geographic region in which they were working. The tracking team used a combination of pay-for-service Internet search engines that pulled from public records and free tools, such as reverse phone searches and white page telephone directories, coupled with extensive telephone networking with family members and contact people to generate leads for the field interviewing team to follow in person when searching for lost respondents. Although telephone and Internet tracking was effective for many cases, a substantial amount of in-person tracking was required to reach the response rate goals. Often, the tracking team would identify a potential address, and the local interviewer would visit the address to find that the person no longer lived there. The local interviewer would then check with neighbors to gather information about where the respondent had moved.

This process sometimes generated additional leads that were then passed back to the tracking team for further Internet searches and telephone work. For example, one fieldworker shared a



tracking story about working with an interviewer, “Tim,” who was traveling in a particular city trying to locate an address where the respondent reportedly owned a house. The first attempt to find the address was unsuccessful because GPS, maps, and the initial in-person visit revealed only the 300-to-500 block of the street name, and the respondent’s address was in the 1400 block. Tim e-mailed the tracker to report the finding, and the tracker double-checked the property records and determined that the respondent’s residence was in the 1400 block of the street in question. The tracker suggested that Tim check with the post office or a municipal office to see if the road continued elsewhere. Another tracker suggested checking with the fire department, which led the interviewer to a new subdivision where he found the address and the respondent. Teamwork and persistence paid off, and the team completed the interview.

Some of the more difficult respondents to locate included youth who had run away or left home without keeping contact with family, respondents intentionally living off the grid, and institutionalized respondents (in nursing homes, detention facilities, and so on). Trust was an issue for some MTO respondents; interviewers required a great deal of persistence to convince some family members that the family was not going to be reported to the authorities. The data collection staff was very successful in getting family members or friends to ask the respondent to call the centralized ISR toll-free line or agree to meet in a public place (such as a library or fast food restaurant) to explain the study and set up an appointment for an interview. The toll-free line forwarded calls directly to the interviewer’s MTO cell phone, which was very helpful because the difficult-to-reach respondents often did not have a phone or had a very limited number of cell-phone minutes. If a call was missed, the respondent may have run out of minutes or would no longer be near a telephone by the time the interviewer was able to return the call. Over the course of the project, the data collection staff tracked respondents who were in jail or detention facilities, noting their release dates and keeping in touch with family members to enlist their assistance in setting up an interview as soon as possible after their release. Some respondents did not speak English or Spanish (the two languages in which the survey instrument was available), requiring translators to help locate and interview family members.

### **Advantages to Real-Time Tracking**

The tracking team used a real-time, web-based sample management system that provided a history of all searches and telephone contacts and enabled trackers to work together to locate particularly difficult-to-find respondents.

At times, the interviewer in the field would visit an address, find it vacant, and telephone the tracker, who would identify another lead or the name of a neighbor while the interviewer was still in the neighborhood. This close collaboration between the tracking and interviewing teams was especially helpful in avoiding the expense of a return trip when travelers were working in non-MTO cities.

Interviewers offered small monetary incentives (finder’s fees) of \$5 to \$10 to contact people who provided the interviewer or tracker with information on the location of a difficult-to-find respondent. ISR added the finder’s fees late in the study, with some limited success. Near the end of the project, finder’s fees rose to \$50 as interviewers worked to find the last of the hardest-to-locate respondents.

Nearly one-half (46 percent) of MTO families required tracking by the tracking team. Even after someone in the family was located, tracking did not stop. A total of 9 percent of sampled individuals (adults and youth) were referred to the tracking team after the family had been located. ISR was able to find and interview 48 percent of respondents referred to the tracking team. The remaining 52 percent were divided among refusals (5 percent), final noninterviews—such as respondents in institutional settings like jails or on military duty abroad—who were not reasonably accessible to the data collection staff (23 percent), and respondents who were randomly excluded from the sample as part of the second-stage subsampling design (24 percent).

### **Strategies To Maximize Respondent Participation**

After locating the respondent, the interviewer's next challenge was to convince each selected participant in the MTO family to agree to complete an interview (and convince parents to give consent for their children to participate). The financial incentive was by far the most effective tool for obtaining respondent participation. Interviewers offered each respondent a \$50 cash payment at the time of the interview and offered adult respondents a \$25 cash payment for agreeing to provide a dried blood spot sample.

ISR implemented additional monetary incentives throughout the data collection period to help meet specific project goals. Consent to audio record the interviews was especially important for MTO, which used audio recordings for quality control checks and for questionnaire items designed to measure the effect of MTO on language patterns. Early in the project period, audio recording consent was lower than projected. In addition to working with interviewers to improve their persuasion skills and increase the number of respondents selected for audio recording, in December 2008, ISR added a \$10 cash incentive for respondents who consented to having their interview recorded. These efforts resulted in an increase in the overall audio consent rate by 10 percentage points, from 74 to 84 percent.

### **End-Game Strategies**

ISR used several strategies at the end of the data collection period in an attempt to obtain the last few interviews needed to raise the response rate to the desired level. We offered additional cash incentives to respondents and interviewers, implemented two-stage subsampling to enable the interviewers to focus limited resources on a smaller number of cases, and extended the study period to allow more time to locate lost respondents and convince reluctant individuals to participate.

As mentioned previously, hiring shortfalls and attrition led to some cities having a smaller data collection staff than planned, with the result being that completion rates (and response rates) varied by site. Rather than implementing end-game strategies at one point in time for the entire sample, ISR added the end-game strategies and incentives on a city-by-city basis.

The survey fielding design employed two-stage subsampling to obtain responses from a representative subsample of hard-to-locate respondents. In stage 1, the ISR data collection staff attempted to contact and interview all the adults and youth in the sample frame. When the response rate at each site reached approximately 75 percent, the team selected a random subset of 35 percent of the remaining cases for more intensive interviewing efforts during stage 2. When calculating the ERR

and analyzing the survey data, respondents interviewed as part of stage 2 received an additional weighting factor so that they represent the other hard-to-reach respondents who were not selected for stage 2 and that ISR did not attempt to interview.

ISR implemented end-game cash incentives in each MTO city based on the completion rate for that city's sample and coupled subsampling with an increase in respondent incentives. ISR increased the interview incentive by \$25 once and then a second time, bringing the incentive offer up to \$100 per interview (plus \$25 for adults for the dried blood spot sample and \$10 for audio recording). When ISR first added end-game incentives in each city, the data collection staff had flexibility in offering the incentives. The interviewer and the team leader discussed each case and decided which cases they were most likely to complete if they offered the additional financial incentive. For the final 2 months of data collection, interview incentives increased to \$200 as the ISR data collection staff attempted to convince the most reluctant and elusive respondents to participate. The final phase of the end-game effort also included two additional options: completing the interview by telephone or completing a shorter version of the survey for a smaller incentive (\$100).

The response rate for respondents to whom interviewers offered an end-game incentive was nearly identical to the response rate for respondents whom interviewers contacted during the end-game period of the study but to whom the interviewers did not offer the end-game incentive. Among respondents who had initially refused to do the interview, 49 percent of those who were offered an end-game incentive and 45 percent of those who were not offered the incentive completed interviews. These findings may reflect the fact that the team leaders used the end-game incentive offers very judiciously, encouraging interviewers to do their best to complete the interview without offering the additional monetary incentive if possible. By the time the end-game incentives were in place, the data collection staff was very experienced and had built up a large toolkit of effective introductions, and their tracking and persuasion skills were well refined.

One aspect of the end-game incentives apparent during the field period was the motivation that they provided to the interviewers. Having something new to say when calling respondents made it easier for interviewers to make additional contact attempts with respondents who had been avoiding them for months. ISR also gave interviewers the flexibility to offer small nonmonetary incentives, such as a small plant or gift, as part of the end-game strategy. These alternative incentives were sometimes effective in helping the interviewer gain access to the respondent's home, thereby enabling the interviewer to pitch the study to a family member or the respondent in an attempt to elicit participation from a reluctant respondent.

### **Strategies To Retain and Motivate Data Collection Staff**

In addition to making efforts to select interviewers who were well suited to the type of interviewing MTO required, ISR focused considerable effort on maintaining staff morale and motivation over the course of a nearly 2-year data collection period. Successful strategies included the continuous efforts of the field supervisory staff, frequent communication among teams and between the ISR central office and the field staff, and a variety of monetary and nonmonetary incentives offered to interviewers and team leaders throughout the data collection period. The direct involvement of the ISR project managers and the NBER research team, along with the importance of the study topic, provided a great deal of motivation to the field staff and helped reduce attrition despite the many

challenges the study presented. The field supervisors of the data collection staff and ISR project managers incorporated continuous interviewer training into all field activities, including standardized training topics in the weekly team meetings, one-on-one coaching with team leaders, working in pairs, and sharing tips and tricks with each other.

One especially motivating technique used during the MTO long-term survey was having the NBER research team meet with the data collection staff. The interviewing and tracking teams met in person in each of the five MTO cities, when possible, and also used telephone conference calls to help contain costs. During the meetings, teams discussed the purpose of the study and progress to date, and, in later months, the NBER research team shared some preliminary demographic information or other simple results with the interviewers. The meetings also included time for discussion: the interviewers provided observations and stories from the field, asked the research team questions, and gathered information that they used in their introductions to help convince reluctant respondents of the importance of the MTO project. The field teams greatly appreciated these meetings, which often resulted in changes to materials or procedures based on interviewers' comments and suggestions. The fact that the research team took the time to meet with—and listen to—the interviewing team was a very positive factor in maintaining high morale and commitment in the field.

ISR offered incentive programs for interviewers at several points in the study. At the end of each calendar year, ISR distributed production bonuses based on productivity and efficiency. In March 2009, ISR held a March Madness competition. The interviewers did not receive any additional compensation, but ISR project managers made a donation to a local food bank in each city based on the number of interviews completed during the month of March. ISR included fun, sports-inspired progress updates in weekly newsletters. This challenge led to a big spike in production and engendered a great deal of team spirit, with the teams adopting names (for example, the New York “Hard Knocks” and the Chicago “Terminators”) and team leaders even engaging in some good-natured competitive trash talk on weekly conference calls. Later in the study period, ISR offered interviewers bonuses for completing high-priority interviews (\$15 for each high-priority interview completed and \$20 if the interview had previously been coded as reluctant). In 2010, as the study was winding down and a very small number of cases remained to work, ISR offered remaining staff a retention bonus for staying on the project and a weekly incentive if they met goals for number of hours worked, followed the work plan for tracking, and made a specified number of contact attempts during the week.

In addition to offering monetary and nonmonetary incentives, ISR provided training opportunities for interviewers and team leaders to help the data collection staff remain productive and effective as the study progressed and as completing interviews became increasingly difficult. Team meetings and newsletter entries provided techniques for convincing reluctant respondents to participate, tracking strategies, and tips for successful interviewing. Newsletters also highlighted success stories, including naming staff to the “Century Club” when they completed 100 interviews and a weekly “kudos” segment recognizing individuals and teams for their efforts. All incentive programs included team leaders. The positive attitude of the team leaders was contagious, and their leadership was a big factor in the success of the field effort. In addition to including team leaders in the incentive programs, ISR developed a professional development series for team leaders, offering seminars on topics such as communication, motivating and getting the best from a team, and work/life integration (recognizing that working from a home office can be very challenging).

## **Coordination Between NBER and ISR**

Throughout the preproduction and data collection periods, the ISR project managers and the field supervisors of the data collection staff worked very closely with the NBER research team, which helped ensure that all parties clearly understood goals, priorities, and the realities of fieldwork. When issues arose, the ISR project managers, field supervisors of the data collection staff, and NBER research team members discussed options and made joint decisions that best met the project needs within the limits and constraints of time, available resources, cost, and quality.

In addition to participating in weekly meetings and countless telephone calls and exchanging e-mails, ISR and NBER developed a wide variety of reports to facilitate the close monitoring and coordination of the field data collection effort. Statistical reports displayed, in tabular and graphical form, information about project cost, data quality, consent rates for components (for example, audio recording or blood spot collection), and completion and response rates. ISR modified the reports over time to better meet the needs of the ISR project managers, field supervisors of the data collection staff, and NBER research team members. This information exchange enabled the data collection staff to identify areas needing additional training or encouragement in the field and to develop programs to meet those needs. The field supervisory staff wrote weekly qualitative reports, providing context and stories from the field that helped to complete the picture of how data collection was progressing.

The ISR project managers, field supervisors of the data collection staff, and NBER research team members carefully monitored completion and response rates across cities, treatment and control groups, and respondent type (adults and youth) throughout the project. Such rates differed for a variety of reasons. Differences in staffing levels led to differential completion rates across sites (for example, the Section 8 group adult sample was released midway through the data collection period [February 2009], when additional funding was obtained) and completion rates naturally differed across the three groups. When differences in completion rates began to appear, ISR added a priority flag to the data collection protocol. This flag was a designation given to selected respondents, with the goal of improving sample balance in completion and response rates. ISR project managers instructed the field team to arrange its work so that team members called and tracked high-priority cases more aggressively, with the goal of completing as many high-priority interviews as possible. As the study progressed and all interviews became more difficult to complete, ISR added a monetary incentive for each high-priority case that was completed as an interview.

The weekly newsletter shared production statistics for the overall project with the data collection staff, and the weekly team calls reviewed team production and efficiency statistics. The NBER research team was fully engaged in monitoring production and worked closely with the ISR project managers and field production managers to review progress, set priorities, identify issues, and develop and implement solutions to problems as they arose.

## **Survey Data Collection Costs**

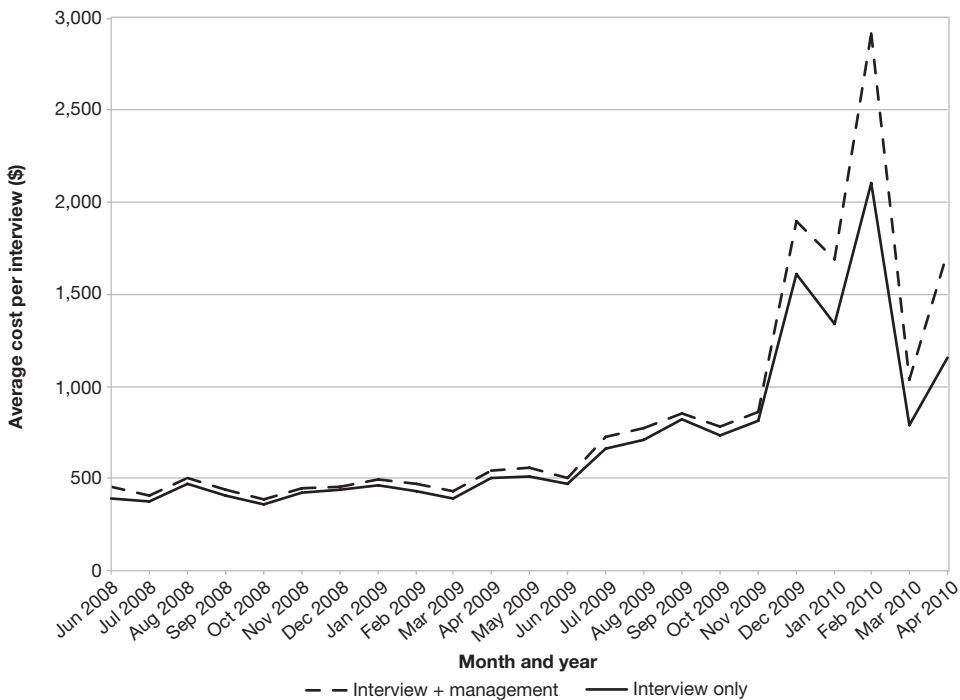
The ISR project managers and NBER research team spent a great deal of time discussing response rate goals and the options for end-game strategies to reach the NBER research team's ambitious goal of 89-to-90-percent ERRs while staying within budget. These debates often took the form of assessing tradeoffs between the high per-interview cost of achieving the last few response rate

percentage points and the benefit to the MTO long-term survey of decreasing nonresponse bias in MTO's estimated effects. When looking at costs, ISR project managers focused on two components: (1) the variable costs associated with obtaining each interview (interviewer hours and nonsalary charges such as mileage, respondent incentive payments, and so on) and (2) fixed costs (for example, for the ISR central office staff to support data collection, project management, creating and checking statistical reports, writing progress reports and weekly memos) that were less dependent on the number of interviews being completed each week. Although the central office staff size decreased along with the field staff size as the study neared completion, the fixed cost per interview increased at the end of the study because those costs were amortized across fewer and fewer completed interviews.

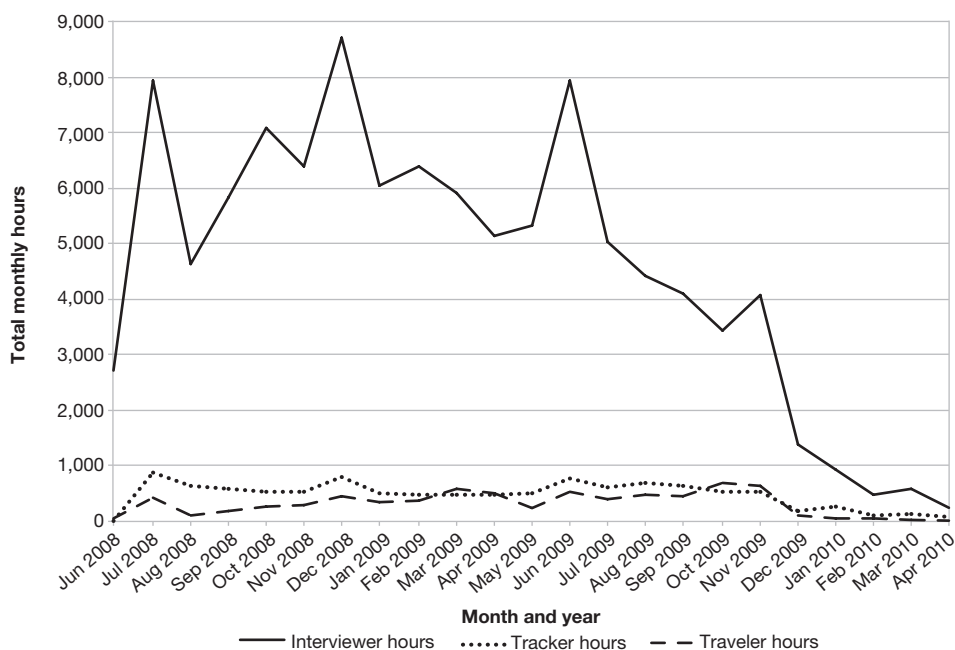
Exhibits 2 and 3 show relatively distinct fluctuation points when survey data collection costs escalated. Exhibit 2 presents average costs per interview throughout MTO survey fielding. Note that these costs are not necessarily real-time costs, because of the delay between the actual interviews and when those costs were entered from an accounting perspective. Nonetheless, average costs per interview were quite steady at about \$470 from July 2008 through July 2009, at which point most of the fresh sample from all three releases had been worked relatively thoroughly. Average interview costs increased substantially in the fall of 2009, to \$802 per interview from August through October 2009 and to \$1,076 per interview for November and December 2009. These costs are not surprising in light of the anticipated work it would take to complete interviews with

**Exhibit 2**

**MTO Long-Term Survey Interview Costs by Month**



MTO = Moving to Opportunity.

**Exhibit 3****MTO Long-Term Survey Interview Hours by Month**

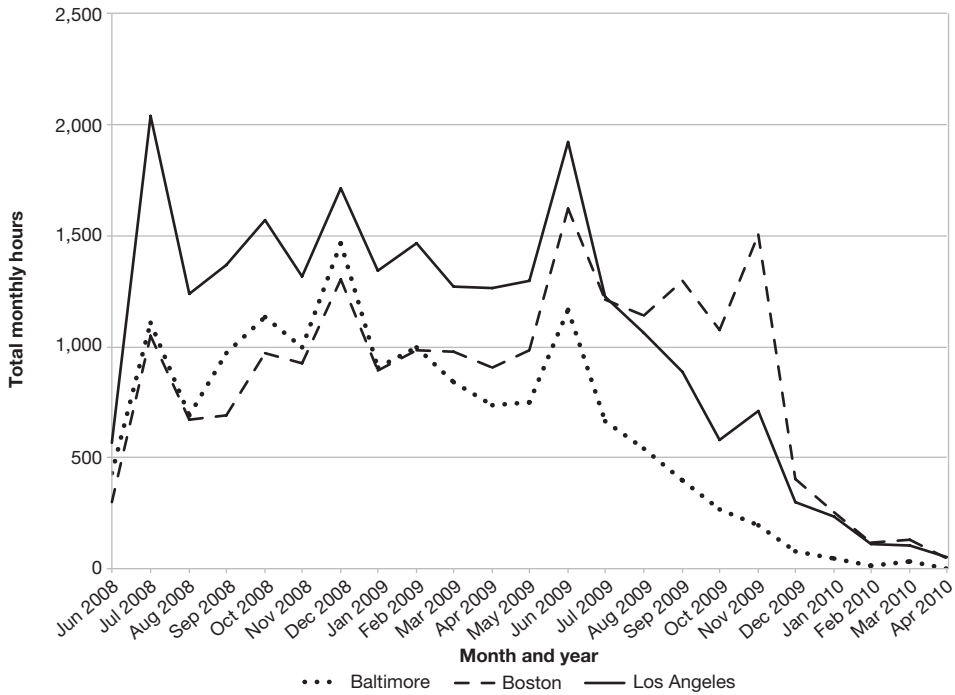
MTO = Moving to Opportunity.

the hardest-to-locate cases. The two-stage subsampling strategy was in place for each site at this point. Average monthly costs per interview escalated starting in December 2009 to about \$1,600. Exhibit 3 presents an alternative but key metric of costs: total interviewer hours and hours spent by travel interviewers (that is, interviewers who were shared across sites or called on duty to travel to zones outside of the immediate area of the five initial MTO sites) and trackers (that is, those ISR staff who located sample using a variety of web, in-person, and alternative techniques, as described previously). Notably, December 2009 represented a key point at which ISR project managers, field supervisors of the data collection staff, and the NBER research team very seriously evaluated the costs and benefits of continuing survey data collection. October 2009 represents another cost flex point, at which time ISR had implemented the two-stage subsampling strategy at many of the sites.

As previously mentioned, the team had to balance meeting the high effective overall response rate with creating a completed survey sample that was balanced by site and by treatment status. Exhibit 4 shows how costs and survey completion rates varied by site. Baltimore had an accelerated survey interview completion rate in terms of reaching ERR targets, whereas Boston and Los Angeles had slower completion rates, in part because of field staff constraints. This exhibit exemplifies the site-based balancing act that ISR project managers and field supervisors of the data collection staff considered in attempting to meet the overall ERR target. This tension in approaches to meet an overall high ERR target, whether via triaging by working the hardest sample for each site or via focusing efforts on one site at a time, was a key balancing act that, as we will discuss in the following section, influenced the study's main findings.

**Exhibit 4**

**MTO Long-Term Survey Interview Hours by Month: Baltimore, Boston, and Los Angeles**



MTO = Moving to Opportunity.

**MTO’s Effects Under Varying ERRs**

The average cost per interview from the beginning of survey fielding (June 2008) through October 2009 was about \$500. As mentioned previously, average interview costs jumped substantially thereafter: to almost \$1,100 in November and December 2009 and to more than \$1,600 from December 2009 through April 2010, when survey fielding ended. The adult survey ERR increased by 5.3 percentage points (285 interviews) between October 2009 and December 2009 and by 3.2 percentage points (97 interviews) between December 2009 and April 2010. This increase roughly translates to \$58,000 per 1-percentage-point gain in ERR (calculated as the cost per interview plus the number of interviews completed) between October and December 2009 and \$49,000 per 1-percentage-point gain in ERR between December 2009 and April 2010. Did the additional dollars spent toward gaining 1 extra percentage point in the ERR measurably or qualitatively alter the main conclusions in the final impacts evaluation (Sanbonmatsu et al., 2011)? We do not conduct a formal cost-benefit analysis but rather employ a simple back-of-the-envelope comparison of survey data collection costs at various time points during survey fielding and with a small number of metrics to attempt to capture the benefit via the MTO demonstration’s contribution to research and policy.



For our back-of-the-envelope analysis, we focused on two representative thresholds that align with observable fluctuations in survey data collection costs: October 2009, when the survey achieved an 81-percent ERR for adults (80 percent for youth), and December 2009, when the survey achieved an 86-percent ERR for adults (85 percent for youth). These ERRs, and the associated dates when they were achieved, also have general appeal. First, we did not want to confound our analyses with the cost efficiencies gained through the two-stage subsampling strategy to allocate more resources per a randomly selected, hard-to-locate case that was triggered at roughly a 75 percent ERR. Second, the ERRs at these cut points generally represent the range of ERRs normally achieved in a wide variety of survey data collection efforts (80 to 90 percent). We use these cut points as simulated dates at which survey fielding ended to construct new samples to reestimate MTO's effects and to examine whether a qualitative difference emerged in three factors: the size of the intention-to-treat (ITT) estimate, the precision of that estimate, and the depiction of the control group. These proposed metrics are of scientific and policy interest; that is, they help inform the following questions: Would our description of the status of the sample have changed had we ended the survey fielding period early? Would our confidence of MTO's effect have changed? Would our interpretation of the program or policy influence on the outcome of interest have changed?

We reanalyzed MTO's effects (for more explanation about the ITT and treatment-on-the-treated [TOT] estimates, see Gennetian et al., 2012; Ludwig, 2012; and Sanbonmatsu et al., 2012) under varying ERR scenarios—those mapped with the December 2009 and October 2009 cut points—in the following manner. First, we replicated the MTO ITT and TOT results for the outcome of interest for the completed MTO long-term survey. Recall that the final ERRs were 90 percent for adults and 89 percent for youth. We then compared the full-sample ITT and TOT estimates with ITT and TOT estimates reanalyzed using the following strategies: (1) using data from the completed pooled sample as of December 31, 2009, reflecting an overall 86-percent ERR for adults (85 percent for youth), and (2) using data from the completed overall sample as of either December 31, 2009 (when some sites, such as Baltimore, achieved something greater than 86-percent ERR), or the date at which the site achieved the equivalent ERR of 86 percent for adults (85 percent for youth). Strategy 2 recognizes the heterogeneity in ERR completion rates by site, whereas strategy 1 is relatively agnostic about site and instead focuses on the pooled ERR. The ERR target for the MTO long-term study was set for the entire MTO survey sample, with an important but secondary target to have a relatively representative sample from each site. In reality, ISR's site-based field staff strategy, coupled with other factors—such as difficulty recruiting or retaining interviewers and the relative ease of finding sample for geographic or comparable reasons—meant that some sites achieved ERR targets faster than others. The variation in site-based survey data completion rates also implies that the date of the last completed interview will vary by site.<sup>7</sup> We replicated strategies 1 and 2 under a

---

<sup>7</sup> Additional analyses that created a sample based on these ERR targets within treatment or control group did not uncover qualitative differences from the final sample results. This result was expected, in part, because by construction, the NBER research team and ISR project managers carefully monitored temporal balance by treatment or control group; that is, that roughly equivalent interviews were being completed for experimental, Section 8, and control group members in any one week or month and, if that was not the case, adjustments were made in real time to achieve this balance by flagging and prioritizing work on selected respondents.

slightly altered assumption of stopping survey fielding as of October 31, 2009, reflecting an overall 81-percent ERR for adults (80 percent for youth). Note in this latter case that Boston, New York, and Los Angeles were just shy of achieving the 81-percent ERR by October 2009.

Exhibits 5, 6, and 7 visually present results for a few of the outcomes and are the focus of our discussion. (Exhibits 8, 9, and 10 at the end of the article provide more detail.) The results shown in the exhibits suggest that the final MTO findings would have differed qualitatively for some of the important outcomes if survey fielding had stopped earlier. For example, if survey fielding had stopped at 81 percent for adults—either pooled or by site—the NBER research team would have falsely concluded that MTO had no effect on two of the four health outcomes. Of the four survey outcome measures for female youth, we would have falsely concluded that MTO had no effect on female youth mental health and that MTO increased female youth idleness (neither employed nor in school).

When examining MTO's effects on neighborhood poverty, exhibit 5 (and the first outcome in the more detailed exhibit 8) suggests little qualitative difference in MTO's effects across the various ERR assumptions, either through the size of the effect, the precision of the effect, or the description of the control group. Turning to MTO effects on other outcomes, exhibits 6 and 7 illustrate a slightly different pattern of results. MTO effects on adult psychological distress are very slightly larger (that is, larger reductions in psychological distress) at an 86-percent ERR. The differences are magnified when comparing MTO's effects for the final sample (90-percent ERR) with an 81-percent ERR. The difference between the analyses at the 86-percent ERR and at the 81-percent ERR is especially pronounced for the within-site ERR adjustments; if survey fielding had stopped when each site reached an 81-percent ERR, MTO's effect on adult psychological distress, at  $-0.128$ , would have been 21 percent larger than the effect for the final sample, at  $-0.107$ . On the other hand, if fielding had stopped at an overall ERR of 81 percent, with variation in ERR by site, MTO effects would have been qualitatively very similar to those estimated at the final 90-percent ERR. Exhibit 7 (and the fifth outcome in the more detailed exhibit 9) suggests a similar, yet even more striking, pattern for female youth: MTO's effects on female youth psychological distress, at  $-0.116$  for the full sample (89-percent ERR), is qualitatively larger and more precisely measured than estimates measured at overall ERRs of 85 percent ( $-0.084$  ITT) or 80 percent ( $-0.050$  ITT).

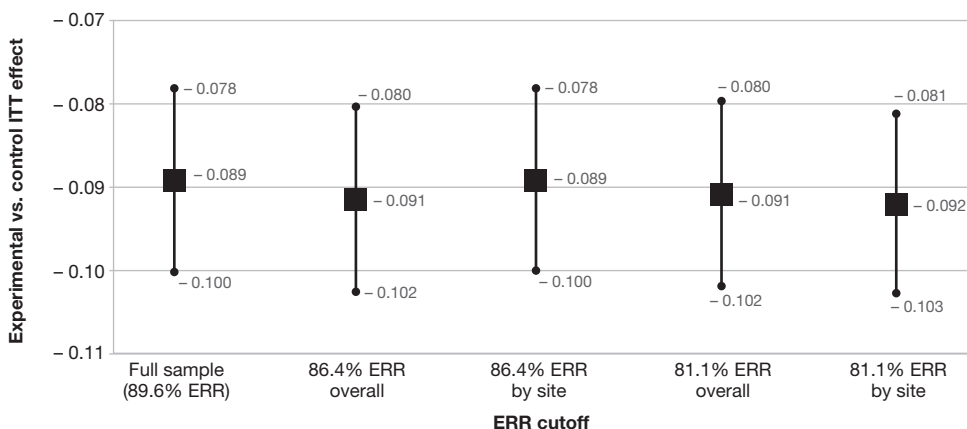
Thus, taking the female youth psychological distress outcome as a starting place, the roughly \$460,000 expended to achieve the last 8.7 percentage points in ERR (between November 2009 and April 2010) translated to a 43-percent difference in the effect estimate.

## **Discussion and Conclusion**

Several strategies contributed to achieving the high response rate goals set for the MTO long-term survey, including selecting and training a data collection team that was well equipped to work in a challenging environment and having staff who understood (and were motivated by) the importance of the MTO demonstration. Starting with a small team and bringing on additional staff after the demonstration started, although not in the original plan, turned out to be very beneficial to a demonstration as complex and difficult as MTO. The close collaboration between the ISR and NBER teams, effective communication between and across the ISR data collection staff, and a solid management structure were also keys to the success of the field effort.

### Exhibit 5

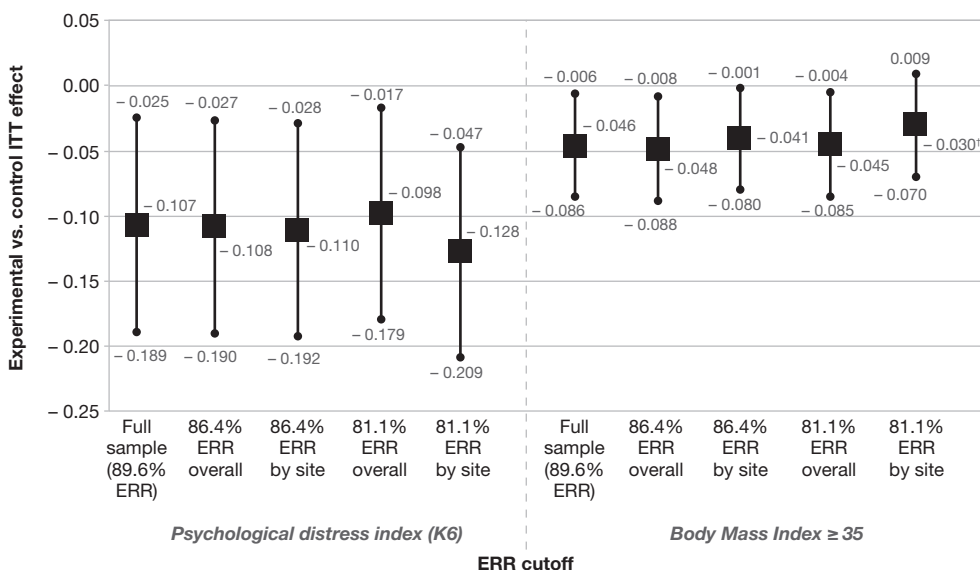
#### ITT Effects on Duration-Weighted Neighborhood Poverty Under Varying Response Rate Assumptions



ERR = effective response rate. ITT = intention to treat.  
 Note: All ITT effects are statistically significant ( $p < .05$ ).

### Exhibit 6

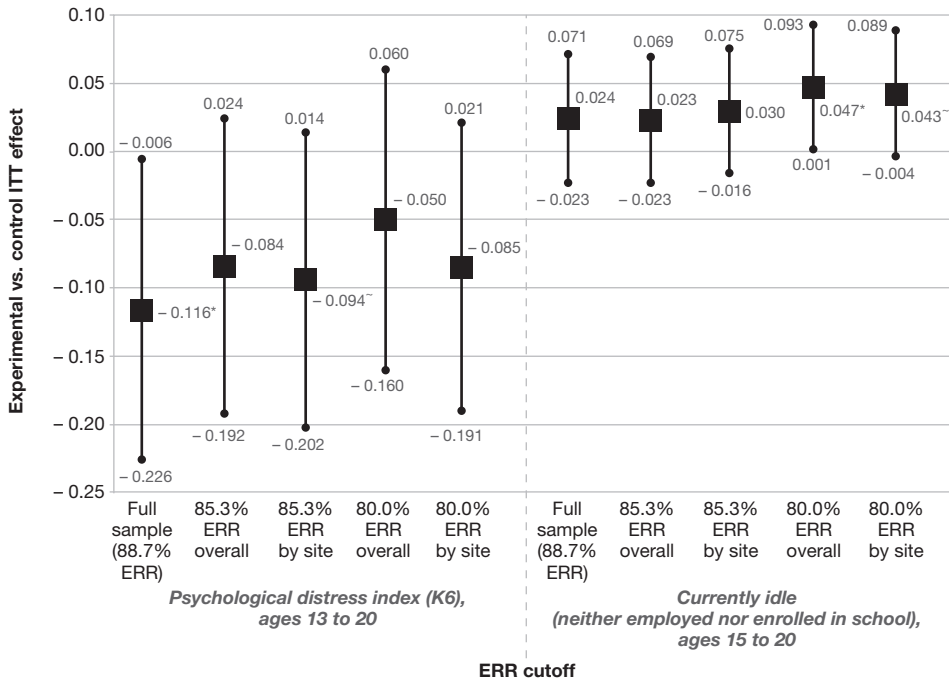
#### ITT Effects on Adult Psychological Distress and Obesity Under Varying Response Rate Assumptions



ERR = effective response rate. ITT = intention to treat.  
 † ITT effect is not statistically significant ( $p < .05$ ).

**Exhibit 7**

**ITT Effects on Female Youth Psychological Distress and Idleness Under Varying Response Rate Assumptions**



ERR = effective response rate. ITT = intention to treat.

\* =  $p < .05$ . ~ =  $p < .10$ .

These strategies all complemented the purer financial incentives that were offered to respondents and interviewers as motivations for behavior. The offer of the end-game incentive did not garner the expected response; the share of respondents who completed an interview was the same between those offered the original incentives and those offered the additional end-game incentives. Keeping the study in the field and refreshing interviewers with new types of incentives to offer to respondents did increase the ERR, because they motivated the interviewers to keep trying to locate and interview elusive and reluctant respondents.

The uneven number and slight variation in the quality of staffing across MTO sites led to differential completion rates and necessitated adjustments to the field procedures (using travelers to supplement local staff), and they also led to incremental implementation of the end-game activities. The strategy of pushing for as-high-as-possible within-site ERRs (as opposed to focusing only on a high overall ERR) certainly put strain on study resources (imagine the cost savings of shutting down one or two sites early, for example) but also added value. The analysis of ITT estimates at different response rates is one marker of potential value, wherein MTO effects on psychological distress of adults and female youth are larger and more precise with the final ERR compared with results that would have been reported had the study stopped when the field reached lower ERRs.

**Exhibit 8**

MTO Effects on Selected Adult Outcomes Under Varying Response Rate Assumptions (1 of 3)

Outcome	Control Mean	Experimental vs. Control		Section 8 vs. Control		Respondents (N)
		ITT	TOT	ITT	TOT	
<b>Share poor: all addresses since random assignment (duration-weighted) [CEN]</b>						
<b>89.6% ERR (no restrictions)</b>	0.396	-0.089* (0.006)	-0.184* (0.012)	-0.069* (0.007)	-0.111* (0.011)	3,270
<b>86.4% ERR</b>						
Overall (December 31, 2009)	0.398	-0.091* (0.006)	-0.185* (0.011)	-0.070* (0.007)	-0.110* (0.011)	3,219
By site	0.395	-0.089* (0.006)	-0.182* (0.011)	-0.068* (0.007)	-0.109* (0.010)	3,213
<b>81.1% ERR</b>						
Overall (October 31, 2009)	0.398	-0.091* (0.006)	-0.184* (0.012)	-0.069* (0.007)	-0.109* (0.011)	3,102
By site	0.397	-0.092* (0.006)	-0.186* (0.011)	-0.073* (0.007)	-0.116* (0.010)	3,118
<b>Felt safe or very safe during the day [SR]</b>						
<b>89.6% ERR (no restrictions)</b>	0.804	0.036* (0.016)	0.074* (0.034)	0.045* (0.021)	0.072* (0.034)	3,262
<b>86.4% ERR</b>						
Overall (December 31, 2009)	0.806	0.036* (0.016)	0.072* (0.033)	0.042* (0.021)	0.067* (0.034)	3,216
By site	0.806	0.036* (0.016)	0.073* (0.033)	0.047* (0.021)	0.075* (0.033)	3,207
<b>81.1% ERR</b>						
Overall (October 31, 2009)	0.812	0.029~ (0.016)	0.060~ (0.032)	0.037~ (0.022)	0.057~ (0.034)	3,099
By site	0.802	0.042* (0.016)	0.085* (0.033)	0.056* (0.021)	0.088* (0.034)	3,115
<b>Rates current housing as excellent or good [SR]</b>						
<b>89.6% ERR (no restrictions)</b>	0.570	0.053* (0.021)	0.109* (0.044)	0.031 (0.029)	0.050 (0.046)	3,267
<b>86.4% ERR</b>						
Overall (December 31, 2009)	0.569	0.044* (0.021)	0.090* (0.043)	0.028 (0.029)	0.045 (0.046)	3,221
By site	0.575	0.046* (0.021)	0.094* (0.043)	0.028 (0.028)	0.044 (0.045)	3,212
<b>81.1% ERR</b>						
Overall (October 31, 2009)	0.577	0.042* (0.021)	0.086* (0.043)	0.028 (0.029)	0.044 (0.045)	3,105
By site	0.577	0.042* (0.021)	0.084* (0.042)	0.032 (0.029)	0.051 (0.045)	3,120

**Exhibit 8**

**MTO Effects on Selected Adult Outcomes Under Varying Response Rate Assumptions (2 of 3)**

Outcome	Control Mean	Experimental vs. Control		Section 8 vs. Control		Respondents (N)
		ITT	TOT	ITT	TOT	
<b>Annual individual earnings (previous calendar year, 2009 dollars) [SR]</b>						
<b>89.6% ERR (no restrictions)</b>	12,288.52	326.93 (583.44)	677.92 (1,209.79)	- 613.60 (807.20)	- 982.43 (1,292.40)	3,141
<b>86.4% ERR</b>						
Overall (December 31, 2009)	12,226.32	116.85 (573.27)	238.58 (1,170.50)	- 655.16 (797.54)	- 1,030.37 (1,254.29)	3,092
By site	12,443.13	284.46 (578.11)	586.33 (1,191.63)	- 1,225.73 (794.67)	- 1,939.20 (1,257.24)	3,089
<b>81.1% ERR</b>						
Overall (October 31, 2009)	12,231.15	147.26 (572.09)	300.22 (1,166.32)	- 438.14 (797.54)	- 680.85 (1,239.34)	2,983
By site	12,258.56	157.97 (571.88)	320.39 (1,159.90)	- 826.60 (774.43)	- 1,294.17 (1,212.49)	3,002
<b>Major depressive disorder with hierarchy (lifetime) [SR]</b>						
<b>89.6% ERR (no restrictions)</b>	0.203	- 0.032~ (0.017)	- 0.066~ (0.035)	- 0.048* (0.021)	- 0.077* (0.034)	3,269
<b>86.4% ERR</b>						
Overall (December 31, 2009)	0.198	- 0.027 (0.016)	- 0.055 (0.033)	- 0.039~ (0.021)	- 0.061~ (0.034)	3,222
By site	0.202	- 0.034* (0.017)	- 0.070* (0.034)	- 0.041~ (0.021)	- 0.065~ (0.034)	3,214
<b>81.1% ERR</b>						
Overall (October 31, 2009)	0.199	- 0.023 (0.017)	- 0.046 (0.034)	- 0.046* (0.021)	- 0.072* (0.033)	3,105
By site	0.205	- 0.036* (0.017)	- 0.073* (0.034)	- 0.038~ (0.022)	- 0.060~ (0.034)	3,121
<b>Psychological distress index (K6) (z-score) [SR]</b>						
<b>89.6% ERR (no restrictions)</b>	0.000	- 0.107* (0.042)	- 0.221* (0.087)	- 0.097~ (0.056)	- 0.156~ (0.091)	3,273
<b>86.4% ERR</b>						
Overall (December 31, 2009)	0.000	- 0.108* (0.042)	- 0.220* (0.084)	- 0.089 (0.056)	- 0.141 (0.089)	3,222
By site	0.000	- 0.110* (0.042)	- 0.226* (0.085)	- 0.109* (0.055)	- 0.173* (0.088)	3,216
<b>81.1% ERR</b>						
Overall (October 31, 2009)	0.000	- 0.098* (0.041)	- 0.199* (0.084)	- 0.096~ (0.057)	- 0.151~ (0.089)	3,105
By site	0.000	- 0.128* (0.041)	- 0.258* (0.084)	- 0.097~ (0.055)	- 0.152~ (0.087)	3,121

**Exhibit 8**

**MTO Effects on Selected Adult Outcomes Under Varying Response Rate Assumptions (3 of 3)**

Outcome	Control Mean	Experimental vs. Control		Section 8 vs. Control		Respondents (N)
		ITT	TOT	ITT	TOT	
<b>Body Mass Index <math>\geq</math> 35 [M, SR]</b>						
<b>89.6% ERR (no restrictions)</b>	0.351	-0.046* (0.020)	-0.095* (0.042)	-0.053* (0.027)	-0.086* (0.043)	3,221
<b>86.4% ERR</b>						
Overall (December 31, 2009)	0.359	-0.048* (0.020)	-0.097* (0.041)	-0.054* (0.027)	-0.085* (0.043)	3,172
By site	0.347	-0.041* (0.020)	-0.083* (0.041)	-0.053* (0.026)	-0.085* (0.042)	3,166
<b>81.1% ERR</b>						
Overall (October 31, 2009)	0.359	-0.045* (0.021)	-0.091* (0.042)	-0.043 (0.027)	-0.068 (0.043)	3,055
By site	0.342	-0.030 (0.020)	-0.061 (0.041)	-0.039 (0.027)	-0.062 (0.042)	3,073
<b>HbA1c test detected diabetes (HbA1c <math>\geq</math> 6.5) [M]</b>						
<b>89.6% ERR (no restrictions)</b>	0.204	-0.052* (0.018)	-0.108* (0.038)	-0.011 (0.024)	-0.017 (0.038)	2,737
<b>86.4% ERR</b>						
Overall (December 31, 2009)	0.203	-0.053* (0.018)	-0.107* (0.037)	-0.013 (0.024)	-0.020 (0.037)	2,700
By site	0.200	-0.046* (0.018)	-0.094* (0.037)	-0.010 (0.023)	-0.016 (0.037)	2,695
<b>81.1% ERR</b>						
Overall (October 31, 2009)	0.199	-0.043* (0.018)	-0.087* (0.037)	0.010 (0.024)	0.016 (0.037)	2,599
By site	0.192	-0.038* (0.018)	-0.078* (0.036)	0.009 (0.023)	0.014 (0.037)	2,613

CEN = 1990 and 2000 census data. ERR = effective response rate. ITT = intention to treat. M = measured. MTO = Moving to Opportunity. SR = self-reported. TOT = treatment on the treated.

\* =  $p < .05$ . ~ =  $p < .10$ .

Notes: Robust standard errors shown in parentheses. The control mean is unadjusted. Unless otherwise indicated, the control mean and effects are expressed as shares of the sample in the category (for example, a control mean of 0.250 for working would indicate that 25 percent of the control group was working). Experimental and Section 8 effects were estimated jointly using an ordinary least squares regression model controlling for baseline covariates, weighted, and clustering on family. Outcomes from the adult survey also control for field release. Census tract characteristics are linearly interpolated from the 1990 and 2000 Decennial Census. Major depressive disorder with hierarchy takes into account the comorbidity of mania and hypomania. The psychological distress index consists of six items (sadness, nervousness, restlessness, hopelessness, feeling that everything is an effort, worthlessness) scaled on a score from 0 (no distress) to 24 (highest distress) and then converted to z-scores using the mean and standard deviation of control group adults. For obesity inputs (height and weight), only a very small percentage of the sample self-reported their height or weight. Body Mass Index is measured as weight in kilograms divided by height in meters squared.

Source: Adult long-term survey

**Exhibit 9**

**MTO Effects on Selected Female Youth Outcomes Under Varying Response Rate Assumptions (1 of 3)**

Outcome	Control Mean	Experimental vs. Control		Section 8 vs. Control		Respondents (N)
		ITT	TOT	ITT	TOT	
<b>Reading assessment score, ages 13–20 [ECLS-K]</b>						
<b>88.7% ERR (no restrictions)</b>	0.000	– 0.020 (0.056)	– 0.040 (0.113)	0.055 (0.063)	0.083 (0.095)	2,286
<b>85.3% ERR</b>						
Overall (December 31, 2009)	0.000	– 0.023 (0.055)	– 0.046 (0.113)	0.062 (0.063)	0.093 (0.095)	2,258
By site	0.000	– 0.033 (0.056)	– 0.068 (0.114)	0.056 (0.063)	0.086 (0.096)	2,252
<b>80.0% ERR</b>						
Overall (October 31, 2009)	0.000	– 0.022 (0.055)	– 0.047 (0.115)	0.071 (0.063)	0.106 (0.093)	2,188
By site	0.000	– 0.016 (0.056)	– 0.033 (0.116)	0.052 (0.064)	0.077 (0.094)	2,196
<b>Math assessment score, ages 13–20 [ECLS-K]</b>						
<b>88.7% ERR (no restrictions)</b>	0.000	– 0.036 (0.060)	– 0.073 (0.121)	– 0.038 (0.067)	– 0.057 (0.101)	2,280
<b>85.3% ERR</b>						
Overall (December 31, 2009)	0.000	– 0.039 (0.060)	– 0.079 (0.121)	– 0.031 (0.067)	– 0.047 (0.102)	2,251
By site	0.000	– 0.051 (0.060)	– 0.104 (0.123)	– 0.026 (0.067)	– 0.040 (0.102)	2,245
<b>80.0% ERR</b>						
Overall (October 31, 2009)	0.000	– 0.046 (0.057)	– 0.097 (0.119)	– 0.006 (0.067)	– 0.009 (0.100)	2,182
By site	0.000	– 0.043 (0.061)	– 0.088 (0.126)	– 0.031 (0.068)	– 0.046 (0.100)	2,190
<b>Currently idle (neither employed nor enrolled in school), ages 15–20 [SR]</b>						
<b>88.7% ERR (no restrictions)</b>	0.194	0.024 (0.024)	0.049 (0.048)	0.031 (0.027)	0.048 (0.043)	1,838
<b>85.3% ERR</b>						
Overall (December 31, 2009)	0.192	0.023 (0.023)	0.047 (0.047)	0.038 (0.027)	0.059 (0.042)	1,807
By site	0.190	0.030 (0.023)	0.060 (0.048)	0.038 (0.027)	0.060 (0.042)	1,806



**Exhibit 9**

MTO Effects on Selected Female Youth Outcomes Under Varying Response Rate Assumptions (2 of 3)

Outcome	Control Mean	Experimental vs. Control		Section 8 vs. Control		Respondents (N)
		ITT	TOT	ITT	TOT	
<b>80.0% ERR</b>						
Overall (October 31, 2009)	0.182	0.047* (0.023)	0.098* (0.049)	0.043 (0.026)	0.065 (0.041)	1,747
By site	0.190	0.043~ (0.024)	0.087~ (0.049)	0.036 (0.026)	0.053 (0.039)	1,753
<b>Risky behavior index, ages 13–20 [SR]</b>						
<b>88.7% ERR (no restrictions)</b>	0.442	-0.027 (0.019)	-0.054 (0.037)	-0.017 (0.020)	-0.026 (0.031)	2,358
<b>85.3% ERR</b>						
Overall (December 31, 2009)	0.443	-0.029 (0.019)	-0.059 (0.038)	-0.017 (0.020)	-0.027 (0.031)	2,326
By site	0.441	-0.028 (0.019)	-0.057 (0.038)	-0.014 (0.020)	-0.023 (0.031)	2,321
<b>80.0% ERR</b>						
Overall (October 31, 2009)	0.440	-0.021 (0.019)	-0.043 (0.039)	-0.005 (0.020)	-0.008 (0.031)	2,256
By site	0.435	-0.020 (0.019)	-0.041 (0.038)	-0.004 (0.020)	-0.006 (0.030)	2,262
<b>Psychological distress index (K6) (z-score), ages 13–20 [SR]</b>						
<b>88.7% ERR (no restrictions)</b>	0.000	-0.116* (0.056)	-0.234* (0.113)	-0.013 (0.065)	-0.020 (0.101)	2,371
<b>85.3% ERR</b>						
Overall (December 31, 2009)	0.000	-0.084 (0.055)	-0.171 (0.112)	0.026 (0.066)	0.040 (0.102)	2,332
By site	0.000	-0.094~ (0.055)	-0.193~ (0.113)	0.021 (0.065)	0.033 (0.103)	2,334
<b>80.0% ERR</b>						
Overall (October 31, 2009)	0.000	-0.050 (0.056)	-0.105 (0.117)	0.055 (0.065)	0.084 (0.099)	2,262
By site	0.000	-0.085 (0.054)	-0.175 (0.111)	0.035 (0.064)	0.052 (0.096)	2,271
<b>Currently good or better health, ages 10–20 [SR]</b>						
<b>88.7% ERR (no restrictions)</b>	0.862	0.003 (0.019)	0.007 (0.038)	0.006 (0.021)	0.010 (0.034)	2,600
<b>85.3% ERR</b>						
Overall (December 31, 2009)	0.866	-0.005 (0.019)	-0.010 (0.038)	0.002 (0.021)	0.003 (0.033)	2,560
By site	0.866	-0.003 (0.019)	-0.006 (0.038)	0.000 (0.021)	0.000 (0.034)	2,561

**Exhibit 9**

**MTO Effects on Selected Female Youth Outcomes Under Varying Response Rate Assumptions (3 of 3)**

Outcome	Control Mean	Experimental vs. Control		Section 8 vs. Control		Respondents (N)
		ITT	TOT	ITT	TOT	
<b>80.0% ERR</b>						
Overall (October 31, 2009)	0.868	-0.011 (0.019)	-0.023 (0.039)	0.005 (0.021)	0.008 (0.032)	2,479
By site	0.865	-0.004 (0.019)	-0.007 (0.039)	0.001 (0.021)	0.001 (0.032)	2,495

ECLS-K = Early Childhood Longitudinal Study-Kindergarten cohort study. ERR = effective response rate. ITT = intention to treat. MTO = Moving to Opportunity. SR = self-reported. TOT = treatment on the treated.

\* =  $p < .05$ . ~ =  $p < .10$ .

Note: Robust standard errors shown in parentheses. The control mean is unadjusted. Unless otherwise indicated, the control mean and effects are expressed as shares of the sample in the category (for example, a control mean of 0.250 for working would indicate that 25 percent of the control group was working). Experimental and Section 8 effects were estimated jointly using an ordinary least squares regression model controlling for baseline covariates, weighted, and clustering on family. Youth and grown children effects by gender were estimated as an interaction with treatment status. Age ranges as of December 2007 are specified for each measure. The reading and math achievement assessment scores are theta scores transformed into z-scores via standardization on the mean and standard deviation for control group youth ages 13 to 20. The risky behavior index is the fraction of four risky behaviors (smoking, alcohol use, marijuana use, and sex) that the youth reports ever having exhibited. The psychological distress index consists of six items (sadness, nervousness, restlessness, hopelessness, feeling that everything is an effort, worthlessness) scaled on a score from 0 (no distress) to 24 (highest distress) and then converted to z-scores using the mean and standard deviation of control group youth. Results reported for the achievement assessment and K6 measures differ slightly from those in Sanbonmatsu et al. (2011) because here, standardization was separate by gender, whereas Sanbonmatsu et al. standardized only on the overall control group mean and standard deviation. The overall (male and female combined) z-score values combine the z-scores by gender and thus are not themselves standardized (the control mean is 0 but the standard deviation is not exactly 1).

Source: Youth long-term survey

**Exhibit 10**

**MTO Effects on Selected Male Youth Outcomes Under Varying Response Rate Assumptions (1 of 3)**

Outcome	Control Mean	Experimental vs. Control		Section 8 vs. Control		Respondents (N)
		ITT	TOT	ITT	TOT	
<b>Reading assessment score, ages 13–20 [ECLS-K]</b>						
<b>88.7% ERR (no restrictions)</b>	0.000	0.027 (0.054)	0.057 (0.115)	0.025 (0.057)	0.035 (0.079)	2,146
<b>85.3% ERR</b>						
Overall (December 31, 2009)	0.000	0.031 (0.055)	0.063 (0.113)	0.033 (0.058)	0.046 (0.081)	2,111
By site	0.000	0.024 (0.055)	0.049 (0.114)	0.036 (0.058)	0.050 (0.081)	2,099
<b>80.0% ERR</b>						
Overall (October 31, 2009)	0.000	0.030 (0.056)	0.063 (0.115)	0.042 (0.059)	0.059 (0.082)	2,027
By site	0.000	0.022 (0.056)	0.045 (0.114)	0.053 (0.060)	0.074 (0.083)	2,036

**Exhibit 10**

MTO Effects on Selected Male Youth Outcomes Under Varying Response Rate Assumptions (2 of 3)

Outcome	Control Mean	Experimental vs. Control		Section 8 vs. Control		Respondents (N)
		ITT	TOT	ITT	TOT	
<b>Math assessment score, ages 13–20 [ECLS-K]</b>						
<b>88.7% ERR (no restrictions)</b>	0.000	– 0.014 (0.056)	– 0.030 (0.119)	0.034 (0.063)	0.046 (0.087)	2,140
<b>85.3% ERR</b>						
Overall (December 31, 2009)	0.000	– 0.012 (0.057)	– 0.024 (0.119)	0.053 (0.063)	0.074 (0.087)	2,105
By site	0.000	– 0.033 (0.057)	– 0.069 (0.118)	0.032 (0.061)	0.044 (0.084)	2,092
<b>80.0% ERR</b>						
Overall (October 31, 2009)	0.000	– 0.020 (0.058)	– 0.040 (0.121)	0.037 (0.062)	0.051 (0.086)	2,021
By site	0.000	– 0.015 (0.058)	– 0.030 (0.118)	0.063 (0.062)	0.087 (0.085)	2,030
<b>Currently idle (neither employed nor enrolled in school), ages 15–20 [SR]</b>						
<b>88.7% ERR (no restrictions)</b>	0.235	– 0.011 (0.027)	– 0.023 (0.058)	0.022 (0.031)	0.032 (0.045)	1,766
<b>85.3% ERR</b>						
Overall (December 31, 2009)	0.235	– 0.017 (0.027)	– 0.036 (0.056)	0.006 (0.030)	0.009 (0.043)	1,724
By site	0.237	– 0.010 (0.027)	– 0.020 (0.056)	– 0.005 (0.029)	– 0.007 (0.042)	1,721
<b>80.0% ERR</b>						
Overall (October 31, 2009)	0.234	– 0.016 (0.027)	– 0.034 (0.056)	0.005 (0.030)	0.007 (0.043)	1,658
By site	0.245	– 0.021 (0.027)	– 0.043 (0.055)	– 0.006 (0.030)	– 0.009 (0.042)	1,664
<b>Risky behavior index, ages 13–20 [SR]</b>						
<b>88.7% ERR (no restrictions)</b>	0.491	0.025 (0.018)	0.053 (0.039)	0.029 (0.020)	0.042 (0.028)	2,265
<b>85.3% ERR</b>						
Overall (December 31, 2009)	0.487	0.028 (0.018)	0.058 (0.038)	0.030 (0.020)	0.042 (0.028)	2,220
By site	0.484	0.020 (0.018)	0.041 (0.038)	0.032 (0.020)	0.046 (0.028)	2,211
<b>80.0% ERR</b>						
Overall (October 31, 2009)	0.482	0.022 (0.019)	0.046 (0.039)	0.030 (0.020)	0.043 (0.028)	2,132
By site	0.485	0.021 (0.018)	0.043 (0.038)	0.030 (0.020)	0.042 (0.028)	2,143

**Exhibit 10**

**MTO Effects on Selected Male Youth Outcomes Under Varying Response Rate Assumptions (3 of 3)**

Outcome	Control Mean	Experimental vs. Control		Section 8 vs. Control		Respondents (N)
		ITT	TOT	ITT	TOT	
<b>Psychological distress index (K6) (z-score), ages 13–20 [SR]</b>						
<b>88.7% ERR (no restrictions)</b>	0.000	0.041 (0.056)	0.088 (0.120)	0.087 (0.063)	0.124 (0.089)	2,273
<b>85.3% ERR</b>						
Overall (December 31, 2009)	0.000	0.047 (0.056)	0.098 (0.117)	0.096 (0.062)	0.135 (0.087)	2,226
By site	0.000	0.056 (0.055)	0.117 (0.116)	0.120~ (0.062)	0.171~ (0.088)	2,218
<b>80.0% ERR</b>						
Overall (October 31, 2009)	0.000	0.075 (0.056)	0.157 (0.117)	0.146* (0.063)	0.206* (0.088)	2,138
By site	0.000	0.048 (0.055)	0.098 (0.114)	0.101 (0.061)	0.143 (0.087)	2,149
<b>Currently good or better health, ages 10–20 [SR]</b>						
<b>88.7% ERR (no restrictions)</b>	0.903	0.006 (0.016)	0.012 (0.035)	-0.007 (0.019)	-0.010 (0.027)	2,500
<b>85.3% ERR</b>						
Overall (December 31, 2009)	0.903	0.009 (0.016)	0.018 (0.034)	-0.006 (0.019)	-0.008 (0.026)	2,453
By site	0.905	0.010 (0.015)	0.021 (0.032)	-0.011 (0.019)	-0.015 (0.027)	2,442
<b>80.0% ERR</b>						
Overall (October 31, 2009)	0.902	0.011 (0.016)	0.023 (0.033)	-0.007 (0.019)	-0.010 (0.027)	2,358
By site	0.906	0.009 (0.016)	0.018 (0.032)	-0.008 (0.019)	-0.012 (0.026)	2,373

ECLS-K = Early Childhood Longitudinal Study-Kindergarten cohort study. ERR = effective response rate. ITT = intention to treat. MTO = Moving to Opportunity. SR = self-reported. TOT = treatment on the treated.

\* =  $p < .05$ . ~ =  $p < .10$ .

Note: Robust standard errors shown in parentheses. The control mean is unadjusted. Unless otherwise indicated, the control mean and effects are expressed as shares of the sample in the category (for example, a control mean of 0.250 for working would indicate that 25 percent of the control group was working). Experimental and Section 8 effects were estimated jointly using an ordinary least squares regression model controlling for baseline covariates, weighted, and clustering on family. Youth and grown children effects by gender were estimated as an interaction with treatment status. Age ranges as of December 2007 are specified for each measure. The reading and math achievement assessment scores are theta scores transformed into z-scores via standardization on the mean and standard deviation of control group youth ages 13 to 20. The risky behavior index is the fraction of four risky behaviors (smoking, alcohol use, marijuana use, and sex) that the youth reports ever having exhibited. The psychological distress index consists of six items (sadness, nervousness, restlessness, hopelessness, feeling that everything is an effort, worthlessness) scaled on a score from 0 (no distress) to 24 (highest distress) and then converted to z-scores using the mean and standard deviation of control group youth. Results reported for the achievement assessment and K6 measures differ slightly from those in Sanbonmatsu et al. (2011) because here, standardization was separate by gender, whereas Sanbonmatsu et al. standardized only on the overall control group mean and standard deviation. The overall (male and female combined) z-score values combine the z-scores by gender and thus are not themselves standardized (the control mean is 0 but the standard deviation is not exactly 1).

Source: Youth long-term survey

## Acknowledgments

The authors thank the Institute for Social Research data collection staff—interviewers, trackers, team leaders, production coordinators, and production managers—for their tireless efforts and dedication to high-quality data collection on the Moving to Opportunity for Fair Housing demonstration.

## Authors

Nancy Gebler is a survey director in the Survey Research Center at the University of Michigan.

Lisa A. Gennetian is a senior research director of economic studies at The Brookings Institution.

Margaret L. Hudson is a survey specialist in the Survey Research Center at the University of Michigan.

Barbara Ward is a survey director in the Survey Research Center at the University of Michigan.

Matthew Sciandra is a research analyst at the National Bureau of Economic Research.

## References

- American Association of Public Opinion Research. 2012. "AAPOR Best Practices." [http://www.aapor.org/Best\\_Practices1.htm](http://www.aapor.org/Best_Practices1.htm) (accessed January 22).
- Gennetian, Lisa A., Matthew Sciandra, Lisa Sanbonmatsu, Jens Ludwig, Lawrence F. Katz, Greg J. Duncan, Jeffrey R. Kling, and Ronald C. Kessler. 2012. "The Long-Term Effects of Moving to Opportunity on Youth Outcomes," *Cityscape* 14 (2): 137–168.
- Gouskova, Elena A. 2008. *PSID Technical Report: The 2005 PSID Transition to Adulthood Supplement Weights*. Ann Arbor, MI: Institute for Social Research.
- Groves, Robert M., Floyd J. Fowler, Mick P. Couper, James M. Lepkowski, and Eleanor Singer. 2004. *Survey Methodology*, 7th ed. New York: John Wiley & Sons.
- Ludwig, Jens. 2012. "Guest Editor's Introduction," *Cityscape* 14 (2): 1–28.
- Macintyre, Sally A., and Anne Ellaway. 2003. "Neighborhoods and Health: An Overview." In *Neighborhoods and Health*, edited by Ichiro Kawachi and Lisa F. Berkman. New York: Oxford University Press: 20–42.
- Orr, Larry, Judith D. Feins, Robin Jacob, Erik Beecroft, Lisa Sanbonmatsu, Lawrence F. Katz, Jeffrey B. Liebman, and Jeffrey R. Kling. 2003. *Moving to Opportunity for Fair Housing Demonstration Program: Interim Impacts Evaluation*. Report prepared by Abt Associates Inc. and the National Bureau of Economic Research for the U.S. Department of Housing and Urban Development, Office of Policy Development and Research. Washington, DC: U.S. Department of Housing and Urban Development.

Sanbonmatsu, Lisa, Jens Ludwig, Larry F. Katz, Lisa A. Gennetian, Greg J. Duncan, Ronald C. Kessler, Emma Adam, Thomas W. McDade, and Stacy Tessler Lindau. 2011. Moving to Opportunity for Fair Housing Demonstration Program: Final Impacts Evaluation. Report prepared by the National Bureau of Economic Research for the U.S. Department of Housing and Urban Development, Office of Policy Development and Research. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

Sanbonmatsu, Lisa, Jordan Marvakov, Nicholas A. Potter, Fanghua Yang, Emma Adam, William J. Congdon, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Jeffrey R. Kling, Ronald C. Kessler, Stacy Tessler Lindau, Jens Ludwig, and Thomas W. McDade. 2012. "The Long-Term Effects of Moving to Opportunity on Adult Health and Economic Self-Sufficiency," *Cityscape* 14 (2): 109–136.

Weiss, Charlene, and Barbara A. Bailar. 2002. "High Response Rates for Low-Income Population In-Person Surveys." In *Studies of Welfare Populations: Data Collection and Research Issues*, edited by Michele Ver Ploeg, Robert A. Moffitt, and Constance F. Citro. Washington, DC: The National Academies Press: 86–104.