Making Connections Sample Design

November 2010

This document serves as an addendum to the "Introduction to Making Connections Data" technical guide. It provides a summary of the sample design employed by Making Connections, focusing on the properties relevant to longitudinal analysis.

I.   Overview of Sample Design

Cross Sectional. The Making Connections sample design is cross-sectional in that in each wave we employ area probability methods to select a sample of addresses to represent each neighborhood. This yields a set of households that are representative of each Making Connections neighborhood at both points in time. All of the Wave 1 addresses (whether an interview was completed there or not) were eligible for selection at Wave 2. Interviewers re-visited these sampled addresses in person or by telephone with the goal of collecting data with whoever was found to be living there. In many instances, the residents have not changed (we call these "stayer" households and give them a **MOVEID** of 0). Other times, new people have moved in (we call these "new households" and give them a **MOVEID** of 4). We also add new addresses at the start of each follow-up effort to include buildings that had been constructed since the previous wave, known as "fresh" housing units (identified with a **MOVEID** of 8).

➲   In a nutshell: in each wave, NORC attempts to conduct an interview at every sampled address, just like all area probability studies.

Longitudinal. The study is also longitudinal in that we are interested in how families *with children* that participated in the baseline are doing. If in Wave 2 we discovered that they had moved away from their baseline address, we followed them for interview at their new location. Households must have had children in the prior wave to be considered a "mover."

➲   In a nutshell: Making Connections also tracks households with children if they move.

II. Screening and Selection Procedures

While the particular method of selecting a respondent and – if there are children in the household, focus child – differs depending on whether or not a given household participated in the past, Making Connections interviewers always begin by enumerating the individuals that live in the household on separate adult and child "Rosters." This approach serves several functions. First, to meet the sampling objectives every adult must have an equal chance to be selected to represent the household. Similarly, in households with children every child must have an equal chance to be selected as the Focus Child.[1] The data from the adult and child rosters can also

---

[1] In the RDD and Wave 1 survey, the interviewer asked a series of questions about just the Focus Child. In Wave 2, interviewers repeated this section for every child in the household.

be used by researchers to investigate how the composition of households impacts family dynamics and experiences in neighborhoods.[2]

After completing the adult and child roster, interviewers proceeded to select a respondent for every household and a Focus Child for every household with children. Because the selection process worked differently in the baseline than in Wave 2, it is described separately for each phase below.

Baseline (RDD and Wave 1)

In the baseline, we employed a semi-random, two-pronged approach to select a respondent to speak on behalf of each household. The key distinction was whether or not there were any children under age 18 living in the household.

Households *with* Children

1. If there was only one child listed on the child roster:
   o That child was selected as the Focus Child.
   o To determine who would be the respondent, interviewers then asked, "Who is the parent or guardian in the household who knows the most about [FOCUS CHILD]?"
2. If there were two or more children listed on the child roster:
   o Interviewers used a Kish table – a random selection tool developed by statistician Leslie Kish – to select one child to be the Focus Child.
   o To determine who would be the respondent, interviewers then asked, "Who is the parent or guardian in the household who knows the most about [FOCUS CHILD]?"

   ➲ In a nutshell: in households with children, the selection of a focus child was random. The selection of the respondent was *not* random.

Households *without* Children

1. If there was only one adult listed on the adult roster:
   o That adult was selected as the respondent.
   o Since there were no children living in the household, no Focus Child was selected.
2. If there were two or more adults listed on the adult roster:
   o Interviewers used a Kish table – a random selections tool developed by statistician Leslie Kish – to select one adult to be the respondent.
   o Since there were no children living in the household, no Focus Child was selected.

   ➲ In a nutshell: in households without children, the selection of the respondent was random.

Wave 2

In Wave 2 interviewers screened all housing units to determine the correct person to interview. The screening procedure was more complex than in the baseline and was driven by two main criteria: 1) whether or not there are children in the household at the time of the Wave 2 interview (the "current household") and 2) whether or not the current household composition had changed since the previous wave. There were six possible procedures, described below.

---

[2] Recall that unlike with Wave 1 and Wave 2, the RDD data were collected at the *county* rather than *neighborhood* level. However, the RDD questionnaire maintained a substantive focus on neighborhoods (in particular, see Segment 2: Neighborhood Connections).

Households *with* Children

1. If there was *no* change between rounds in the composition of *children* (in other words, no new children were added to the child roster)
   - o Interviewers kept the same Focus Child as was selected in the previous wave.
   - o If the previous wave respondent still lived in the household, he/she was re-selected to be the respondent. Otherwise, interviewers asked, "Who is the parent or guardian in the household who knows the most about [FOCUS CHILD]?" and proceeded to choose a new respondent.

2. Otherwise[3]:
   - o Interviewers used the Kish table to randomly select a Focus Child.
   - o To determine who would be the respondent, interviewers then asked, "Who is the parent or guardian in the household who knows the most about [FOCUS CHILD]?"
     - ▪ Note that the selected adult who may or may not have been the same person chosen as the respondent in the previous wave.

➲ In a nutshell: in households with children, the selection of a focus child was sometimes random, sometimes not. The selection of the respondent was *not* random.

Households *without* Children

1. If there was *no* change between rounds in the composition of *adults* (in other words, no new adults were added to the adult roster):
   - o We re-interviewed the same respondent as the last time.
   - o Since there were no children living in the household, no Focus Child was selected.
2. Otherwise:[4]
   - o An adult was selected using the Kish table to be the respondent.
   - o Since there were no children living in the household, no Focus Child was selected.

➲ In a nutshell: in households without children, the selection of the respondent was sometimes random, sometimes not.

Because the screening and selection procedure for Wave 2 was complex, we provide some general rules and hints in the box on the following page.

---

[3] This includes (a) cases freshly added to the Wave 2 sample (**MOVEID**=8), (b) "stayers" (**MOVEID**=0) who were not successfully interviewed in the past because they were deemed out of scope (OOS), refused to participate, or could not be reached, and (c) stayers who *were* interviewed in the previous wave and were found to include a new child(ren) in Wave 2.
[4] Again, this includes cases freshly added to the Wave 2 sample (**MOVEID**=8) as well as "stayers" (**MOVEID**=0) who were not successfully interviewed in the past because they were deemed out of scope (OOS), refused to participate, or could not be reached.

GENERAL RULES FOR WAVE 2
1. As in the baseline, interviewers aimed to collect data at *all* sampled addresses.
2. New to Wave 2, we started following "movers":
   a. We followed and attempted to complete a full interview with an adult living with any previous wave focus child who was still under age 18 and had moved out of a sampled address. These "mover" interviews are identified in the data sets with a **MOVEID** of 2.
   b. We also followed households with an aged-out focus child (one that had turned 18 years of age or older) to ask a couple questions about the aged-out focus child. The resulting abbreviated interviews are identified in the data sets with a **MOVEID** of 6.

Hints
1. Remember that only Wave 1 households that included children were eligible to become "movers."
2. By "move" we do not mean that an entire household relocated from their Wave 1 housing unit as an entity. Rather, for Wave 1 households with children, we define "move" as the focal child residing in a different housing unit at Wave 2. For Wave 1 households without children, we base this definition on the Wave 1 respondent.

III. Properties of the Data

As a result of the sampling design outlined above, the Making Connections data can provide[5] the following:

1. <u>For cross-sectional analysis:</u>
   a. **A sample of households that is representative of a given target neighborhood** (Wave 1, Wave 2). All Wave 1 cases can be included because they were randomly sampled from a list of USPS addresses. When using the Wave 2 data, filter where MOVEID = 0, 4, or 8. This will include households who were interviewed at the same address in Wave 2 ("stayers"), new households who moved into those addresses sampled in Wave 1 ("new households"), and inhabitants of newly-sampled addresses ("fresh cases"). Excluded will be Wave 1 focus children who moved to a new address but remained under age 18 in Wave 2 ("movers") and Wave 1 focus children who moved to a new address and also turned age 18 by Wave 2 ("aged-out movers"). Note that the definitions of neighborhoods are specific to this survey and do not always align with official boundaries, municipal or otherwise.
   b. **A sample of households that is representative of a given county** (RDD).

2. <u>For longitudinal analysis:</u>
   a. **A panel of households with children.** The panel of households interviewed at both Wave 1 and Wave 2 is representative of households with children who resided in the Making Connections neighborhood at Wave 1.
   b. **A panel of adult-only households who did not move.** Households that did not include children in Wave 1 were re-interviewed in Wave 2 provided that they remained at the same address. Note that the respondent may have changed if new adults were found in the Wave 2 household.

Conversely, the Making Connections data *cannot* provide the following:

---

[5] With proper filtering of cases

1. **A representative sample of adults**. Respondents within households were not selected at random if the household contained children, due to the study's emphasis on obtaining information regarding children. So, one should make the distinction between questions concerning households vs. those concerning adults, as the latter would be non-random.
2. **A nationally representative sample of individuals, households, neighborhoods, or cities**. While findings from the ten Making Connections sites are indicative of some experiences and processes prevalent in many urban areas, the data should not be used to make inferences at the national level.
3. **A representative panel of households without children who moved.** We do not follow households without children that moved outside their original housing unit between waves.
4. **A panel representative of counties**. County-level data were collected in the baseline in a random digit dial (RDD) telephone survey, but those households were *not* included in the subsequent follow-up efforts. Therefore, the RDD data are not appropriate for longitudinal analysis. They should be used in isolation or in conjunction with the Wave 1 data[6] to investigate attitudes and experiences at one period in time.

IV. Preparing the Data for Longitudinal Analysis

Researchers will notice that **PRV_SUID** in the Wave 2 files is not unique. Some cases from Wave 1 "spawned" a new interview in Wave 2, thereby creating two records out of one (both sharing the same value at **PRV_SUID**). This happens when the focus child from the previous round moves away from the original interview address. If they were still under age 18 at Wave 2, we would have followed the child to his/her new *address* and also interviewed the new *people* living at the child's old address. Again, this is one of the unique properties of Making Connections: the sample design is both cross-sectional and longitudinal.

To eliminate duplication in the Wave 1 data you are going to have to refine the join properties and probably drop some Wave 2 cases. Here are suggestions:

1. If you want to do longitudinal analysis at the <u>respondent (individual)</u> level, you can filter the Wave 2 data using the variable **R_SAME_AS_WAVE1**. Link the Wave 1 data to the Wave 2 data only where **R_SAME_AS_WAVE1** in the latter = 1 (yes).
2. To do analysis at the <u>household</u> level, use **MOVEID** in the wave 2 data to keep only those Wave 2 records where **MOVEID**=0 (same people, same address) or 2 (same HH with children, new address). Depending on which variables you plan to include in your analysis, you may also want to isolate just those cases with the same respondent in both waves (see #1 directly above). This is particularly important when examining the data from attitudinal questions or those that ask specifically about the experiences of the respondent.
3. To do analysis at the <u>address</u> level (regardless of whether or not the same people lived there across waves), include Wave 2 cases where **MOVEID**=0 (same people, same address) or 4 (new people, same address).

Additional applications of *Making Connections* data may be possible. If you have questions about the sample design or data collection methods, please contact Kate Bachtell, the associate project director at NORC, at bachtell-kate@norc.org.

---

[6] The RDD data allow for a comparison between conditions in the targeted Making Connections neighborhoods and conditions in their host counties.