



## Tracking in Longitudinal Household Surveys

**Firman Witoelar**

**July 6, 2011**

Sample attrition is one of the inherent challenges faced by any longitudinal household survey. Among surveys in developing countries, mobility accounts for most of the attrition. Recent experience in the collection of longitudinal household and individual data around the world has shown that it is feasible to track respondents, i.e. to follow respondents who moved from the location where they were first interviewed. Tracking individuals and households can entail significant costs and may require specific focus by the organization conducting the survey. Drawing from experiences from previous and ongoing surveys, the paper presents recommendations on methodology for successfully implementing tracking in panel household surveys.

---

The author thanks Kathleen Beegle, Gero Carletto, Stefan Dercon, Andrew Dillon, Kristen Himelein, Randal Kuhn, and Talip Kilic, for comments on an earlier version of the paper. This paper was supported by the Knowledge for Change Trust Fund and the LSMS-ISA program in the World Bank Development Research Group. The views expressed are those of the author alone and do not reflect the position of the World Bank. All errors are the sole responsibility of the author.

## 1. Introduction

Sample attrition is one of the inherent challenges faced by any longitudinal household survey. Attrition in household surveys can come from deaths, refusals, inability to locate the baseline dwelling (due to insufficient information about the location), and migration.<sup>1,2</sup> Hill (2004) reviews attrition in thirteen panel studies in developing countries and concludes that mobility accounts for the majority of survey attrition. The issue of attrition and mobility is particularly relevant for longitudinal household surveys that aim to generate nationally representative data with a focus on agriculture and rural development, such as the surveys conducted under the World Bank's Living Standards Measurement Study-Integrated Surveys on Agriculture (LSMS-ISA) project.<sup>3</sup> To gain a better understanding of the spatial development and the linkages between farm and non-farm activities in the countries under the study, understanding the geographical mobility of the survey respondents becomes important. Glewwe and Jacoby (2000) discuss the advantages and disadvantages of collecting panel data in LSMS-type surveys, and provide practical advice on how to collect the data. This paper partially builds on their work, focusing on how to follow respondents when they have moved from their baseline location.

In a longitudinal household survey, the survey planners conducting follow-up interviews need to make decisions about whether to follow households and individuals who have left their baseline location, and to what degree with respect to time, distance, and cost. Recent experience in the collection of longitudinal household and individual data around the world has shown that it is feasible to follow respondents who moved from the location when they were first interviewed. In some surveys, tracked respondents - the number of respondents found by tracking who would otherwise been lost - can amount to 45% of the sample (Hill 2004).

Following the movers and interviewing them not only helps minimize the attrition rate, but also addresses potential selectivity biases in non-random attrition. People who move are unlikely to be a random subset of the baseline respondents. They are likely to have certain characteristics that differ from those who remain in the baseline location. Analysis of panel respondents using only those who remain in their baseline dwelling or in their baseline community will likely suffer from selectivity biases caused by the non-random attrition. The seminal work on panel attrition by Fitzgerald et al (1998) provides some methods to assess the extent of attrition bias by looking at selection on observed

---

<sup>1</sup> Attrition in developed countries is likely to be due to relatively more refusals. Beckett et al (1988) and Fitzgerald et al (1998) provide discussion on attrition in the context of the Panel Study of Income Dynamics in the US. Uhrig (2008) discusses attrition in the context of the British Household Panel Survey.

<sup>2</sup> Inability to locate the baseline dwelling can occur in areas where formal addresses are not used (such as rural areas) and when the information collected at baseline is insufficient. This may happen when there was not an original intent to conduct a panel of the households, and so the baseline data on location is scant. GPS data on location at baseline can be used to facilitate back-tracking to the baseline dwelling in subsequent rounds of the survey.

<sup>3</sup> See the Living Standards Measurement Study-Integrated Surveys on Agriculture (LSMS-ISA) website for a description of the project: <http://www.worldbank.org/lSMS-ISA>

characteristics.<sup>4</sup> When the observed baseline characteristics of movers and stayers are different, there are strong incentives to track movers. However, even if the observed baseline traits are not different, researchers still need to be concerned about analyzing a panel of households restricted to non-movers. A recent study by Thomas et al (2010) on attrition in the Indonesia Family Life Survey shows that failure to follow movers in longitudinal surveys results in higher rates of attrition, attrition selected on observed characteristics, and attrition on *unobserved* characteristics as well, which could complicate inferences.<sup>5</sup> Beegle et al. (2008) show that movers and stayers had similar economic status at baseline, but those who have moved out of the study area in Tanzania experienced much higher consumption growth than those who stayed. Using only the sample of those who remained, based on the assumption that the lack of observable differences at baseline would not impact sample selection, would in fact seriously underestimate poverty changes in the panel of households. These findings emphasize that it is questionable to argue that attrition through migration will not introduce biases when the baseline characteristics of movers and stayers are not different.<sup>6</sup>

In the event that an entire household has left the baseline dwelling, an alternative to trying to locate the baseline household (or some of its previous members) is to replace that household with the new occupants of the baseline (original) dwelling.<sup>7</sup> However, to claim that this maintains the representativeness of the sample implies a very strong assumption that the new entrants and those who left are interchangeable. On the issue of interchangeability, Rosenzweig (2003) warns against using survey rules that condition on household residence (the dwelling itself). In a study using two rounds of the Bangladesh Nutrition Surveys (1982 and 2002), he demonstrates the biases that arose because the survey re-interviewed only those individuals that were still living in the dwelling originally surveyed.

Recent field experience and research findings thus suggest that longitudinal household surveys should seriously consider tracking respondents who move, and in particular not only track those who move from their baseline dwelling within the community, but also those who move outside the baseline locality. With those analytical advantages in mind, it is important to note that tracking individuals and households can entail significant costs and may require specific focus by the organization conducting the

---

<sup>4</sup> Fitzgerald et al (1998) distinguishes between attrition that is selected on observable and unobservable characteristics. While the paper suggests some tests to look at potential attrition biases based on observables, the authors recognize that assessing the extent of selection on unobservables is much more difficult. The tests proposed by Fitzgerald et al (1998) have been used in many studies. For example, Alderman et al (2001) in their analysis on panel data in developing countries use the tests and do not find significant effect of attrition on estimated coefficients in some multivariate models of household and individual outcomes.

<sup>5</sup> Individual fixed effects could potentially remove these unobserved characteristics, but the IFLS study suggests that there are changes in the unobserved characteristics after the baseline.

<sup>6</sup> It is important to note that while tracking may reduce attrition and significantly mitigate bias induced by non-random migration, the selection bias may still be present, as tracked respondents are likely to be a non-random sample of movers. There is a literature on how to deal with survey non-response by reweighting the data: see for example Korinek, Mistiaen, and Ravallion (2007).

<sup>7</sup> Glewwe and Jacoby (2000) discuss the disadvantages of following the dwellings, which include the need to regularly update the list of dwellings in the area to keep the survey representative of the population.

survey.<sup>8</sup> Drawing on experiences from recent and ongoing panel surveys, this paper will lay out some recommendations on methods for the successful implementation of tracking in panel household surveys.

## 2. The movers: households or individuals?

In longitudinal surveys, attempting to follow households and individuals who moved from original dwellings may lead to significant complexities both at the conceptual and practical levels.<sup>9</sup> Since households split up and regroup in different ways, survey planners need to decide what it means to re-interview “the original household” when not all members still reside together. One household interviewed at baseline can become two in the follow-up round. Children move out, couples divorce, and so on. Questions then arise as to which of these households at follow-up should be regarded as the “successor” of the original household. At the conceptual level, defining which of the households is regarded to be the successor of the original household is problematic, because i) household membership and composition are constantly in flux, ii) household structure and boundaries are not always clear, and iii) the definition of “household” as a unit of analysis will likely depend on the nature of the research question posed by the survey planners. Moreover, household structure and composition are often endogenous to other decisions made by the household or household members. For the purposes of this paper, the panel household survey should be viewed as a survey that follows *individuals* over time, even where, in practice, this entails following individuals no longer residing together as a household after the baseline survey. From a tracking perspective, once an individual is found, the household is then built around the individual.<sup>10</sup>

Analytically, researchers may be interested in following the different units of analysis over time: individuals, households, and dynasties.<sup>11</sup> Some research questions can only be addressed by tracking individuals over time, such education, health, or labor market outcomes, even though information about changes at the household or even dynasty level will also be utilized.<sup>12</sup> Other research questions can

---

<sup>8</sup> In some settings, there may be very few migrants (especially depending on the amount of time between rounds of the survey). The decision about whether to track should be informed by the extent of migration expected. Developing complicated tracking protocols and establishing a separate tracking team may not be worth the costs and effort if, for example, there are very few movers. It is usually not possible to know the extent of tracking before field work starts, but using other data and knowledge of the setting can help inform appropriate planning for tracking.

<sup>9</sup> Partly because of this conceptual problem, Glewwe and Jacoby (2000) argue that LSMS-type surveys should not be based on a sample design that tries to follow households over time. They recommend following either dwellings or individuals over time. In this paper, we argue for following individuals over time.

<sup>10</sup> Once the household is built around the tracked individuals, it is up to the survey planners to determine whether the household is the “same” as in previous round.

<sup>11</sup> The term ‘dynasty’ in this paper refers to the set of households in subsequent rounds of the survey whose members belonged to the same household in the baseline survey. In the literature, a dynasty is sometimes used interchangeably with ‘extended family’ or ‘linked households’.

<sup>12</sup> See Witoelar (2005) and Beegle et al (2008) for examples of analyses that link households originating from the same baseline household. Witoelar uses the constructed panel of extended families to look at consumption risk-

only be answered if the survey tracks the households, and yet other questions can be answered better if the survey tracks dynasties. For example, the issue of asset ownership and its change over time across the three types of unit of analysis would be a matter of particular policy importance.

At the practical level, rules need to be made at the start of field work about which new households to interview in the event that the baseline household is not intact (i.e., where not all members still co-reside, excepting for deaths of members). One option is to define households found in the baseline dwelling as the original households and to define new households in other dwellings where baseline members are found as “split-off” households. In some cases, all members may have moved from the baseline dwelling to new dwellings, and it is not clear which of these households should be thought of as the “original” household. For example, a couple may have divorced and now live in two different dwellings: which household is the “original” one? In the fieldwork, this difficulty often manifests in the assigning of household identification (ID). Existing surveys differ in how they assign household ID. The Indonesia Family Life Survey (IFLS), for example, uses what is called the first-contact-rule, which assigns baseline household ID to whichever household of the set of households with baseline members was contacted first during the resurvey (Frankenberg and Thomas, 2000). In the previous example, the household of the ex-spouse who was first contacted would be labeled the baseline household. The Kagera Health and Development Survey, KHDS 2004, in contrast, assigns a new household ID to each of the households originating from the same baseline household in 1991. Section 9 will provide specific recommendations on how to assign household and individual ID, but the discussion above highlights the problem of conceptualizing a “panel household” (also see Appendix Figure for some examples).

### **3. Establishing tracking rules: who to track?**

The rules on who to track and who to interview should be based on the overall survey design and the aims of the study. One principal consideration will be the representativeness of the panel sample. The baseline survey is typically representative of either the whole population or a particular subset of the population. As respondents move and leave their baseline household, the remaining sample of non-movers will become less representative of the baseline, particularly since migration is not likely to be random. Depending on the objectives of the study, the survey could decide to follow all individuals to ensure that the next round of the survey is representative of the baseline of individuals, or the survey could decide to follow individuals with certain characteristics, as will be discussed below.<sup>13</sup>

Once the survey planners have decided who the tracking targets are, a set of tracking rules to be used in fieldwork should be established. The rules must be simple enough for interviewers to follow during

---

sharing within extended households in Indonesia. Beegle et al use a similar approach to address potential mortality bias when looking at consumption growth among households in Tanzania.

<sup>13</sup> For each case, there are statistical methods and weighting techniques that are proposed so that the sample can regain some of its baseline representativeness. For example, the IFLS produces a set of alternative weights for analysts to consider (Strauss et al 2008). However, the use of weights to recover sample representativeness is not uncontroversial. Technical discussions on representativeness and sampling weights are beyond the scope of this paper. For a more formal discussion on these issues, see Lynn (2009).

fieldwork. Where there are cases that are not covered by pre-established rules, interviewers must refer these back to the survey headquarters for decision to avoid unnecessary tracking effort and cost.

In thinking about tracking, we make a distinction between *tracking targets* and *interview targets*.

**Tracking targets** are those individuals who when moved, should be tracked, and when found, should be interviewed. **Interview targets** are those who need to be interviewed. Due to cost and other limitations, not all interview targets are tracking targets. Interview targets may include respondents from the previous round, who, when moved, need not be tracked. In addition, interview targets may also include individuals who were never before included in the survey and are new entrants to the new households. For example, one could decide that when the tracking target is found in a new household, his/her new spouse and children in that new household will be interview targets, even though the spouse and children were not part of the baseline survey.

There are several potential ways to determine who should be the tracking targets, which include:

- **Tracking targets based on the relationship to household head in the baseline.** In a longitudinal household survey designed to analyze the socioeconomic dimensions of families and household members over time, it may make sense to prioritize tracking the movers who have strong familial links to the other baseline household members. For example, some surveys include relatives and servants as members of the household. When some of the household members move, the head of the baseline household, his/her spouse, and/or his/her children could be given higher priority, and others with more tenuous links (such as distant relatives of the head or servants) given less priority. This consideration will vary across surveys depending how households are defined and on the purposes of the survey.
- **Tracking targets based on certain demographic characteristics.** For some purposes, demographic variables such as age and sex could be used as factors to determine the tracking targets. Depending on specific goals of the survey, individuals with certain characteristics can be targeted. For example, The Young Lives Project, a multi-country study of childhood poverty, follows two cohorts of 12,000 children in Ethiopia, Peru, India, and Vietnam over 15 years. The project makes an effort to follow the children when they change location (Leon an Dercon, 2008). For a study on marriage transition in Malawi, Beegle, Özler, and Poulin (2010) track young adults over time, but do not follow other household members who are not the focus of the study. If the survey did not track these young adults who moved outside their baseline village, it would have missed many of the respondents that were the focus of the study, and estimates based solely on the respondents who stayed would likely have been biased. Another example would be a longitudinal survey that focuses on international migrants. In most surveys, respondents who move out of country become “out of scope” and are dropped from the sample. However, in some contexts, such as the Mexican Family Life Surveys (MxFLS), this group could be the one of the groups on which the study is focused. In this case, the tracking of these respondents can be prioritized. Note that tracking outside the country will add significant complexity and entail much higher costs.

- **Tracking targets based on the individual interview in the baseline survey.** In the baseline survey, it is possible that not all household members will be individually interviewed (for individual information related to, for example, labor, health and education) either by design or otherwise. Some individuals might have refused, some might have not been found, and others might have only been interviewed through proxy interviews. In this case, representativeness of the initial sample may be affected, although this can be corrected for by weighting. Survey planners may decide not to track these household members in subsequent rounds.

In some cases, the entire household may have moved intact to another location. In that case, if all of the tracking targets are found in the new dwelling, then the tracking for the baseline household is complete. However, the survey team will not necessarily know before they track the new dwelling that all members still co-reside in the new dwelling. Thus, it is necessary to have rules about who to track before the field work begins.

The rules about which respondents to track have important ramifications for sample representativeness. Weights that take this non-random selection of tracking targets into account may have to be constructed, even with tracking, since usually not all movers can be found.

#### **4. How far to track?**

There are several aspects to tracking that need to be considered before field work for the follow-up round begins. This includes rules about how far individuals will be tracked if they have moved, in order to manage the costs of tracking efforts. We examine two areas of focus: administrative boundaries and geographic boundaries or distances.

One way to limit how far interviewers should follow tracking targets is to establish the administrative units outside which the targets will not be followed by any team. The administrative units can be districts, counties, provinces, or countries, depending on the context. For example, the IFLS limits tracking to movers who move to one of the 13 provinces covered by the IFLS baseline. The second wave of the Matlab Health and Socioeconomic Surveys (MHSS2) plans to follow migrant households within Bangladesh and about 400 households overseas (see the Appendix Table). There will be different cost, logistical, and perhaps most difficult, administrative implications associated with domestic vs. international tracking that the survey planner must carefully consider. One practical reason for limiting the administrative areas is related to rules concerning obtaining permission from local leaders in each survey area (for example, through official letters announcing the survey, etc). For this reason, the survey organizers should consider obtaining permissions *before* the start of the fieldwork from local leaders/government in the areas that are not part of the baseline sample but where the respondent may likely move. Getting permission to track beyond country borders will be significantly more complicated and often infeasible due to budgetary implications.

One point to consider when limiting tracking to specific administrative areas or units is that on occasion a group of households or individuals may have moved to a locality that sits just outside the

administrative boundaries. The cost in terms of money and time to track them is therefore only marginally higher than the costs of tracking individuals in the localities inside the boundaries. In this case it may be advisable for the survey organizers to require the interviewers to track them.

Another way to limit tracking is to establish geographical/natural boundaries beyond which tracking targets will not be followed. For example, the survey may have a rule to not follow the individuals if they move across to another island or over a mountain range even though the new location is still within the same administrative boundaries. The limit could also be based on distance or length of travel. The survey must set specific guidelines on the maximum distance (e.g., 20 kilometers) or the maximum travel time (e.g., one hour, one-way) under which the interviewers need to find and visit the movers.

It is important to note that the rules of limiting tracking to individuals who are relatively easier to track (be it in the survey districts or nearby geographically) may result in losing individuals who are very different from others. While the cost of interviewing the difficult-to-find individuals may be higher, the scientific cost of dropping them may also be significant (Thomas et al 2010). One way to compromise on this issue is to track everyone who moved relatively close to the EA, and then make strong efforts to track a subset, drawn randomly, of those who have moved further. This was done, for example, in the IFLS. In the ongoing Uganda National Panel Survey (conducted under the LSMS-ISA project), prior to the fieldwork, 20 percent of households found at the baseline enumeration areas in 2008/2009 were randomly chosen to be tracking targets. All household members who move from those selected households are tracked to their new location when they move. In practice, tracking will always be limited by costs and by the fatigue of the fieldwork, so these randomizing efforts can help to minimize bias of the movers who are re-interviewed.

## 5. Timing of the survey

Timing of the survey can play a key role in increasing the success of re-contacting individuals who have migrated from their baseline communities. For instance, in several countries there is a particular time of the year when out-migrants return to their baseline or parental villages to visit their relatives. In China and Vietnam, the period around the new Lunar year is the time that many migrants return to their villages. The period around the Eid (a Muslim holiday) is the time when many Indonesians return to their villages to visit their relatives, leaving big cities deserted. Some surveys have taken advantage of using special holidays to re-contact the out-migrants and collect information from them. Done correctly, this can avoid having to track movers.

There are potential complications that need to be considered with regards to conducting the survey during these important holidays. *First*, the time period is, by definition, a special period. This may have implications for some of the data that are collected. For example, consumption expenditures during this period may be higher or the pattern may be very different from other times of the year. This may also be true for many other variables that could be of interest. *Second*, respondents may decline to participate if they think that by visiting the households during this period, the survey is too intrusive, or they simply do not want the limited time they have with their relatives to be spent answering the

survey. *Third*, finding interviewers who are willing to do the interview during this period can be difficult and/or the salary to pay the interviewers may need to be higher for them to be willing to work during the holiday.

Since the holidays are typically short and respondents may not want to spend the holidays answering questionnaires, an alternative use of the interviewers' and respondents' time during this period is as an opportunity to establish contact with the movers, collect information about their current residence, and based on that information, to only conduct interviews if the respondents live beyond the reach of the resurvey. Instead, if the respondents live nearby, the contact established during the holidays can be used to schedule interviews that will be conducted in the respondents' residence after the holidays.

## 6. Organizing the tracking operation

In this section we discuss how to organize tracking in a panel survey, including structuring the phases of field work, defining the role of a tracking manager, and establishing tracking teams. Survey fieldwork is often done using mobile teams, where each team includes a supervisor, several enumerators, and in some cases a driver, local guide, or translator. Each team would typically have pre-determined routes and time schedules, and would be responsible for a set of households or individuals residing in a particular geographic area (often clustered in enumeration areas, or EAs).<sup>14</sup> Tracking will add many complexities in the operations of these teams. It will mean that the routes and time schedule become more flexible, under coordination from the survey organizers. It will also mean that teams may have to interact with each other, under the coordination and management of the supervisors. This section will provide some recommendations as to how tracking activities should be managed.

### Tracking phases

One way to organize tracking is to divide the field work into three different phases: a pre-fieldwork phase, a main fieldwork phase, and a post-fieldwork tracking phase.

The pre-fieldwork tracking phase is an effort before the full start of the survey to pilot materials for tracking and to potentially gauge the extent of migration in the panel. It entails visiting at least a subsample of survey EAs and collecting minimal information. Collecting information about movers prior to the main fieldwork could be useful for several reasons. *First*, it may help assess the extent of migration among the sample, which will help the survey planners in making decisions about how much tracking should be done. *Second*, pre-field work tracking can help survey planners to improve tracking instruments and protocols for the main fieldwork and post-fieldwork tracking. *Third*, experience gained during pre-fieldwork tracking can be useful to both survey planners and interviewers. Finally, it may help inform the survey planners whether it is at all feasible to conduct another round of survey. The advantages of this strategy includes lower cost, improved questionnaire design, and potentially lower

---

<sup>14</sup> Alternatively, some surveys have resident enumerators who may reside in an enumeration area to conduct all household surveys over some duration. When tracking is involved, however, the need for resident enumerators will be outweighed by the need to have enumerators that are highly mobile.

attrition since the survey has made contact with the households. This pre-fieldwork tracking activity could also include creating maps or geo-referencing the households and the communities to be used during the main fieldwork.

During the main field work, baseline communities (EAs) would be revisited by the survey team. Tracking targets who moved locally (*local movers*, who reside within or in a nearby community) would also be interviewed by the same teams in the respective EAs. These tracking cases would be interviewed while the team is in the EA, extending the team stay in the community if necessary.

In some cases, a tracking target may be identified as having moved to the area which is near the survey route of another field team. In this case, coordination between teams is necessary and it is crucial to have a tracking manager (discussed below) at the headquarters coordinating this work. This type of tracking can be quite important since it turns what is essentially long distance tracking (when the tracking targets are followed by the same team who interviewed the person at baseline) into a lower cost and local tracking effort, done by the team responsible for the survey area to where the tracking target has moved.

One implication of tracking being done during the main field work is that all interviewers must be trained in how to track. Training sessions (in class and in field practices, if applicable) should include practice on how to follow tracking protocols.<sup>15</sup>

For cases of movers who are no longer in or near the baseline and community, and not near another EA in the survey (*long distance movers*), tracking can be done during a tracking phase after the main field work. In this case, the cost for long distance tracking will be high, so tracking needs to be prioritized. In most cases, these movers will not reside near many other movers, so the unit cost of each tracked case is higher than interviews during the main field work. For the tracking phase, smaller teams might need to be configured. Some cost-saving strategies may include sending smaller teams to track. For example, it would be very costly to send an entire team to find one person in a very distant location. From the cost perspective, the decision on whether to pursue these far-away movers can be based on how clustered the individuals to be tracked are. For example, in tracking individuals who moved from Albania to Greece as part of the 2005 Albania LSMS, the Institute of Statistics (INSTAT) sent Greece-based teams to particular areas only if the number of movers in those areas was above a certain threshold, particularly when that involved traveling long distances by sea.<sup>16</sup> When possible, the team should

---

<sup>15</sup> In preparation of the Tanzania National Panel Survey 2010/2011 (part of the LSMS-ISA Project), a field practice for tracking was conducted in both rural and urban area. The field practice took advantage of the availability of unused household listings that were collected by the Tanzania National Bureau of Statistics two years earlier. Using the household listings from a small number of EAs, both rural and urban, a team consisting of survey organizers and field supervisors revisited these households, employed the tracking protocols, and proceeded to track and visit the local movers. A small number of long distance movers were also contacted by phone, and some were visited. The field practice helped train the supervisors on how to follow tracking protocols and at the same time provide inputs on how the tracking instruments and protocols could be improved.

<sup>16</sup> Based on personal communication with Calogero Carletto of the World Bank.

contact the respondents in advance to confirm their locations and make appointments for the interview if necessary.

During the tracking phase, successfully tracked respondents should be reported to the survey headquarters continuously so the team can prioritize the duration of the tracking phase. The role of a tracking manager in monitoring and coordinating tracking activities is crucial.

Deviations from the original timeline of interviews, by the virtue of separating main field work from tracking, may have broader analytical implications (e.g. seasonality) for the measurement of welfare outcomes that the researcher may need to account for ex-post. For example, if all long-distance tracking is done after the main-fieldwork, then the researcher needs to take into account that not only are the data collected during tracking from areas that could be significantly different from the main fieldwork areas, the data are also collected in different times of the year.

Finally, depending on the settings, it may be useful to begin the main survey areas in EAs where more individuals are expected to move out (e.g. rural areas) rather than in areas that are known to be major migrant destinations (e.g. urban or industrial centers). With this strategy, there may be less need for return visits to enumeration areas for tracking purposes. This strategy was done in the Malaysia Family Life Survey 1988 where most migrants are known to move from rural to urban areas. Again, for reasons discussed above, this strategy may not be desirable for household surveys that have an agricultural component where seasonality is an issue.

#### **Tracking management team/tracking manager**

It is absolutely necessary to have a tracking management team or a tracking manager responsible for overseeing the tracking aspects of the survey. Experience has shown that it is crucial to have one person or a small team assigned to the task of monitoring attrition during the main field work and organizing tracking activities.<sup>17</sup>

A tracking management team or tracking manager in the survey headquarters is responsible for several areas. These include:

- **Managing the information system** on data about tracking targets. During the fieldwork the survey teams will send information that they collect about the tracking targets. Many times the teams may also request information about the targets or confirmation about the information that they possess from the headquarters. A successful tracking operation relies on almost real-time information flows between the survey teams and the headquarters. The tracking manager designs and maintains the management information system for tracking. For pencil-and-paper based surveys, one challenge is to make sure that all of the information is recorded

---

<sup>17</sup> Two basic requirements that a tracking manager must possess: "...1) knowledge of the dimensions of migration in the research setting and 2) a vested interest in the quality of follow up that will demand attention to tracking...." (electronic mail correspondence with Randall Kuhn, MHSS). The IFLS2-4 also employed tracking managers to work full time on managing tracking operations.

electronically so they can be communicated efficiently.

- **Solving conflicting information.** One team may send multiple sets of information, and although it may happen more rarely, information may come from different teams about the same person. Information coming from different informants could be conflicting, and the likelihood of this is higher when the information is coming from different teams in different locations. This could happen simply because one informant has better information than the other. The relationship of the informant to the tracking targets, as well as the frequency and intensity of contact between the informant and the mover (which should be included as questions in the Tracking Forms), could serve as guidelines about which informant is likely to have better information.
- **Evaluating quality of tracking information.** The tracking manager will evaluate the quality of tracking information, and which tracking cases have higher probability of being successfully tracked based on the information available. For example, a tracking case with a complete street address, phone numbers, and the sketches of maps from multiple informants that are consistent with each other is more likely to be solved than a tracking case with only a basic location description from a single informant. Cases with higher quality of information may be prioritized. Phone numbers that are collected from the informant, which should have been verified by the enumerators, should be verified again.
- **Organizing and prioritizing the tracking cases by region.** For each region, cases that are very likely, somewhat likely, or unlikely to be solved should be clearly indicated. Although in practice it may be difficult, it may be important not to share this information with all interviewers if there is a concern that the efforts of the interviewer will be affected for cases that are already classified as “unlikely to be solved”. Use telephone and other communication tools to call respondents and make appointments if necessary.
- **Evaluating the trade-off between cases.** If the survey is seeking to get as many re-contacts as possible, teams may prioritize cases that are very likely to be solved first. Given a fixed budget and time constraints, there will always be a trade off, for example, between whether to follow and interview five persons who live close by or two who live far away. This is again where the issue of the interchangeability of respondents comes into play. Is it more valuable to interview five households who look very similar to each other, or to make the effort to follow two households who moved far away? This relates closely to the survey design and thus decisions need to be made from the top, not by the fieldworkers.
- **Assigning tracking cases to teams.** Having determined which cases to prioritize, the tracking manager must assign the cases to the teams in such a way as to optimize the cost and time of the fieldwork. Some cases may need to be bundled so as to save cost. For example, instead of sending a tracking team to village X each time that a tracking target is known to have moved there, it may be useful to wait until several targets have moved there or close by before sending

the interviewers to track the cases in one trip. Bundling the tracking cases requires some judgment on the side of the tracking manager, since the number of cases that can be bundled is unknown until the end of field work. Good coordination with the team is essential. For example, the interviewers may be about to leave an area when information about new tracking cases in that area is received by the headquarters.

- **Record-keeping on tracking.** It is important to have good record-keeping on the tracking cases, including the contact and interview completion rates, and on which team is responsible for which cases. Tracking targets should be sorted by areas, teams, and if applicable, by difficulty or completeness of tracking information. The contact and completion rate should be continuously updated so the tracking manager can immediately identify the enumeration areas and teams that are lagging behind. Transferring tracking cases between teams should be done with the knowledge and permission of the tracking manager, so that the manager can always follow up the case to the appropriate team.
- **Monitor on the extent of moves/migration.** The survey team will only know the extent of the actual migration after a few weeks of fieldwork. Field teams may find out that many of the respondents have moved such that they need fewer days than planned in the EA to complete interviews. If this happens, the survey team and the tracking manager should re-evaluate the plan for the fieldwork, the routes, and the time schedule. For example, in the Uganda National Panel Survey 2009/2010, an evaluation was conducted after the first month of fieldwork to assess the extent of mobility, as there was not much information available before the fieldwork began to predict the extent of tracking that would be needed.
- **Pre-fieldwork tracking.** Even if a full-scale pre-fieldwork tracking cannot be conducted, efforts should be made to conduct at least a smaller version of it. A pre-survey tracking survey could be conducted in several EAs. A small number of EAs are chosen, and baseline households are revisited. Information is then collected on the number of individuals in the baseline sample from the community who have moved, and where they may now reside. This enables survey planners to have some idea of the extent of tracking that needs to be done and what rules need to be established. Pre-fieldwork tracking can also be conducted as a pilot for tracking, from which improvements can be made to the protocols as well as tracking instruments.

In addition to having a fully dedicated tracking manager, it is also important that the supervisors as well as the enumerators be trained in the importance and full scientific benefits of tracking. For this purpose, tracking activities should be well covered in the training of the supervisors and enumerators, and should be viewed as an integral part of the whole fieldwork. Experience suggests that when tracking is introduced as an additional activity to the fieldwork, fieldworkers tend to be less willing to put extra effort and prioritize away from tracking.

### **Dedicated tracking team**

In many contexts, the survey planners may have a good idea where people in the population are likely to migrate to, if they migrate at all. Urban centers and industrial regions, for example, are typical destinations of rural migrants. If this is the case, it may be useful to have a dedicated tracking team working there. This team can set up their base camp in the same area where migrants usually stay. The team can start working as soon as tracking forms start to come in during the main fieldwork; they do not have to wait until after the main fieldwork ends.

The members of such a tracking team should be highly mobile and persistent. Tracking requires high mobility. Pursuing leads about one tracking case can send the interviewer from one end of town to another. Where possible, the use of motorcycle is recommended as it allows one to navigate crowded roads faster. Interviewers should be familiar with the area or should have a very good sense of direction, and should be able to orient themselves quickly to new areas. Additionally, it is of course essential that they be good interviewers, as the efforts spent on finding the tracking target will be wasted if the interviewers cannot administer the interviews properly. In IFLS, interviewers who were involved in long-distance tracking activities were chosen from the best interviewers, since tracking is seen as a particularly challenging task. One alternative to having a team of excellent interviewers who are also good at finding addresses is to recruit local fieldworkers with the sole job of finding and confirming addresses.

## **7. Protocols for tracking**

### **Identifying the dwelling and the individuals**

During the main fieldwork, whenever the interviewers are unable to locate a tracking target in the baseline dwelling, tracking protocols are implemented. The tracking protocol starts with the interviewer completing a Tracking Form with the information necessary to find the movers. However, before concluding that the household (or individuals) has moved, the team should be sure they are at the correct dwelling. Street address, notes, sketches about the location, and if applicable, photos collected in the baseline or previous rounds should be consulted. Also, when applicable, GPS back-tracking can be utilized to ensure the dwelling is the right one. This is less problematic if only some individuals have moved, since the interviewers can match the current (remaining) residents with the household rosters from the previous survey.<sup>18</sup>

Having ensured that they are at the correct dwelling, the interviewers will have to identify whether the entire household (all baseline members) or any household member has moved. Interviewers should come to the household prepared with previous roster(s) of the households. The pre-loaded/pre-printed rosters should list everyone who was previously on the household roster. Basic information such as sex, age, date of birth, and relationship to the head should be included. This is another place where the use

---

<sup>18</sup> This paper discusses the types of basic information from the baseline survey that should be pre-printed/pre-loaded from the perspective of tracking. It does not review other information from the baseline that can also be pre-printed/pre-loaded, such as assets owned and their disposition, enterprises, and land holdings.

of computer-assisted personal interviewing (CAPI) may have some advantages over pencil-and-paper interviews. Logistically, preparing pre-printed rosters on paper, distributing them to the right teams, and making sure the teams use the correct pre-printed forms in the field takes a significant amount of work and training. The use of CAPI can make the process simpler and the logistics much more manageable.<sup>19</sup> Recently, some surveys have also employed the use of photographs of household members to help interviewers match the person to the roster. In the future, biometric data will likely be used to establish identity.

The interviewer should go over the list of each person in the roster to verify whether the individual still resides in the dwelling. For each person listed in the roster, there will typically be at least three possible statuses: 1) present in the household, 2) not present in the household, 3) died. If the individual is present, other information from the pre-printed/pre-loaded roster can be used to verify that he/she indeed is the correct person (photographs are particularly useful here). Individuals no longer present are either movers or are deceased. For deceased persons, one can consider a separate mortality questionnaire. For the movers, Tracking Forms need to be completed.

### **Tracking Form for Movers**

In order to complete a Tracking Form for a mover, the interviewer should try to obtain multiple, reliable informants. It is therefore very useful to have a list of potential informants in the baseline survey who could provide information of the respondents' whereabouts should the respondents decide to move in the future. The information about the potential informants should also be in possession of the interviewer during the resurvey. Collecting information on potential informants should become a standard in longitudinal household surveys, and the information should be collected in each round, not just the baseline.

The importance of having multiple informants who can provide information on where the tracking targets have moved cannot be overstated. Often only partial information about the tracking target is known to a particular informant; having multiple informants may help in providing a complete description of the tracking targets' whereabouts. Some informants may give conflicting information and the conflicting information should be recorded if it cannot be resolved.

In cases where the entire households have moved, the potential informants include neighbors, close relatives living nearby or afar, former employers, former teachers/school employees, or village officials. Some surveys ask the respondents in the baseline survey to list people who could provide information about the respondents' whereabouts should the respondents decide to move in the future. This information should be in possession of the interviewers in the following round. The interviewers should start the tracking process by finding these informants identified at baseline.

---

<sup>19</sup> Among the LSMS-ISA surveys, the ongoing Uganda National Panel Survey is the first to use CAPI for its tracking activities and household interviews.

In cases where only some members have moved, then the household head or other remaining household members found in the baseline dwelling are primary candidates to be the informants. Even then, it is still useful to collect information from other potential informants, as listed above. It is recommended that the tracking form contains a list of these categories of informants and requires the interviewer to assess the availability and completeness of information obtained from each of the categories.

In addition, information from other part of the questionnaires collected in the baseline or previous rounds can also be helpful. For example, if information was collected about the tracking target's place of work or schools in the baseline, interviewers may be able to go to these former places of work or schools to obtain information.

The Tracking Form collects at least three types of information: data on the informants, information on the tracking target's current location, and, possibly, some current socioeconomic information about the tracking target (see the Appendix Table for links to some of the longitudinal household survey official websites, many of which provide links to their survey instruments including the Tracking Forms).

Interviewers should collect information from the informants and record this on a well-designed tracking form, completed for each tracking target. The Tracking Forms should include information about the informants themselves. This includes home or cell phone numbers, address/route, description of location, email address, work place, and work place phone numbers. This is very useful if during the process of tracking the interviewers need more information from the same informant. It is recommended to call the phone numbers that were provided to make sure that the numbers are valid. The interviewers can ask for collaboration from the informant or the household to establish contact with the movers. The Tracking Form should also record the informant's relationship to the target respondent, the frequency and intensity of contact between the informant and the target respondent, and the last time of contact between them. This information will be useful should there be conflicting information from different informants.

With respect to information on the tracking target's current location, there are a number of pieces of information that can be collected about the tracking target which may be useful in locating that individual. This includes:

- New street address of the tracking targets.
- Telephone numbers. In many countries, people have multiple phone numbers and some change their telephone numbers frequently. It is important to obtain all the phone numbers that are available and to try to verify them. One of the best options is to ask a household member to call from their own cell phones since some people may be reluctant to answer calls from unknown numbers. The interviewer could also ask the household member to introduce the interviewer to the mover or even help the interviewer schedule an interview with the mover. Another option is to call the phone numbers given by the informants immediately, even in the presence of the

informants, to make sure that the numbers that are provided are valid. In verifying the phone numbers, care should be taken by the interviewer so as not to offend the informant by suggesting that the enumerator does not trust the informant.

- Route to reach the new address. The route needed to be taken to reach the tracking targets may be available from the informant who has visited the tracking target, even though the informant does not have the target's specific street address. In this case, the interviewers should ask the informant how he reached the target's address, e.g., which public transportation was used, travel time, etc. Informant should be asked to provide a sketch of a map if possible. Landmarks such as bridges, markets, churches, mosques, important intersections, or natural landmarks often provide crucial information and must be recorded.
- New/old work place: the name of the company/employer/place of work that the tracking target is now working for, or worked for before he/she moved, or the name of the tracking target's place of work/company if he/she is self-employed, along with its address and phone numbers. Information about the place of work can be used to find the tracking target or to find another informant that may know of the target's whereabouts. If the respondent/mover is a mobile trader, questions should be asked about the area or location of the markets served by the mover.
- Information about anyone in the tracking target's new location or his/her work place that can be helpful. Informants who have some information about the tracking target's new location or new place of work may also have some information about another person in the new location or new place of work that may be able to provide better information about the target's whereabouts.
- Information about anyone in the baseline area who could also act as informants.

The quality of information on the tracking form from the informants should be assessed. Not all information will be useful for locating the respondent to be tracked, especially if it is partial or incomplete. In these cases, the interviewer may need to find a second informant.

If some household members were found, but others need to be tracked, then in addition to collecting the information necessary for tracking on the Tracking Form, basic information about the respondents who moved could also be collected from the members who are found. This will ensure that at least some information is available if the tracked persons are not located for an interview.

The information collected about the tracking target can include:

- Date the person last resided with the tracking target

- Reason they left (work, marriage, school, etc.) If the move is thought to be because of marriage, then the information could serve as a lower bound for the potential new entrants at the new residence/household.
- Highest education
- Working status, marriage status, and other key information

When an entire household moves, one possible alternative to completing tracking forms is to administer part of the household roster questions in addition to the tracking form to the informants. Again, this will at least ensure that some basic information about the respondents is obtained should they not be found. This task will certainly place more of a burden on the informants and interviewers but may prove to be very valuable.

## **8. Finding the targets**

In addition to using information from the tracking forms, below are several steps that can be conducted to make tracking more successful, some of which have been mentioned above.

**Official/unofficial records.** When available, official records such as birth, marriage or death records, either in the baseline location or in the destination, may be helpful in providing information about the target's whereabouts. Hill (2004) lists official and unofficial records that have been used to track respondents, including: adoption records, phone books, voter/census lists, school records, church membership or organization membership records, and electricity meter readers.

**Information from previous rounds.** As discussed above, information from other modules in the survey that were not specifically collected for tracking purposes may turn out to be very helpful for tracking. For example, information about place of work in the employment module, school in the education module, and health providers in the health utilization module, are all useful sources of information about both the tracking targets and the potential informants. This strategy has been employed in other surveys such as the IFLS. With CAPI technology pre-loading this information into tracking rosters or making it easily accessible for tracking purposes, this will not be difficult.

**Use of mobile phones.** The survey should take advantage of the widespread availability of mobile phones. Mobile phones can be used to contact the tracking targets, confirm their locations, and to make appointments for interviews. They can also be used to contact the informants to confirm or clarify information that was provided earlier. Mobile phone numbers should therefore be collected from the informants whenever possible. This also implies that the teams must be equipped with mobile phones and the budget has to be allocated to buy mobile airtime.

**Photographs.** Photographs of respondents can be utilized to increase re-contact rates in the next survey rounds. Photographs of respondents are taken with the respondents holding a paper/card with their person ID (and household ID) on it. During the resurvey, the interviewers can verify that the individuals

they are visiting are indeed the target respondents. Another option is to take photographs for the entire households with information about the HHID and each person ID noted in the picture. This is to make sure that during the re-survey, the interviewers can refer back to the picture to avoid missing a respondent. Care must be taken to protect the privacy of the respondents.

**The use of CAPI.** The use of CAPI, which has the potential to offer advantages over pencil-and-paper interviews (see for example, de Weerdt, 2009), should be extended to tracking. As discussed in Section 7, CAPI can help in identifying tracking targets and in “building” the households around the individuals. Another possible use for CAPI in tracking includes, as discussed above, pre-loading information not specifically collected for tracking from previous rounds and making it accessible to interviewers. CAPI can make use of digital photographs that have a clear advantage over printed photographs. CAPI can also help save the time needed to collect the tracking information and send them to the headquarters. Since timing and coordination of different teams is critical in organizing tracking activities, CAPI could make a significant difference in the success of tracking.

**GPS backtracking.** When applicable, GIS data can be used to track households that are hard to locate. Some GPS devices have a feature that can provide direction when the user enters the coordinates. The GPS devices with this feature allow the user to enter a specific set of coordinates. Once entered, the device can be told to point the user in the direction and to calculate the distance to the destination. The device will help the user to reach the general area of the household, from where it is usually fairly easy to ask the precise location of the household.

**Gifts/incentives to households.** Providing households with gifts or incentives can increase response rates. A study on the use of respondent incentives on longitudinal surveys in developed countries suggests that incentives tend to be more effective if they are set in advance of the interview, and are in cash rather than in-kind. Higher incentives are more effective and incentives often work better for those least likely to respond (Laurie and Lynn, 2009). However, in developing countries, the effects of incentives can vary. Field experience in some surveys in developing countries suggests that the gifts can be nominal, or does not have to even nearly compensate the respondents for the time spent for the interview. Also, to protect the privacy of respondents, care should be taken that in-kind gifts cannot be used to identify the respondents (e.g. hats or t-shirts that bear the name of the survey should be avoided). Lastly, the ethical implications of providing gifts or incentives to participate should be carefully considered by the organizers of the survey.

**Incentives to tracking teams.** Survey organizers can also consider providing incentives to the tracking teams. The incentives can be based on the number of tracking cases solved and they may also vary according to the difficulties of each case. Interviewers must also be informed about the ethical and legal implications of tracking (Hill 2004).

## **9. Conducting interviews with the tracked respondents**

Once a tracked respondent is found and the interview can be conducted, there are several steps to ensure that the survey will be able to link individuals and households across several survey rounds correctly.

### **Assigning household ID.**

The first step to take when the tracking target is found is to assign a household ID. Rules to assign unique household ID should be established beforehand. Individual or person IDs, usually resulting from the combination of household ID and roster from baseline, should also be maintained across rounds.

Household and individual ID has to be unique in a given round. In panel surveys with tracking, new households are added to the sample when the tracked respondents are found and interviewed. This means that some rules have to be established on how to assign household ID to the new households. There are several options that have been used in other surveys.

- The first option is to assign the baseline household ID to one of the households, and assign new household IDs to the split-off households. The IFLS is one of the surveys that use this rule. The baseline household is given the baseline household ID which is the same as in the previous survey round, but with an additional “00” at the end. Households that split from this baseline household are given a new household ID which share the same household ID with the baseline households, except that for the last 2 digit “00” is replaced with a number identifying the split-off (see example 1 in the Appendix Figure).
- The second option is to assign a new household ID to all households in each round by simply adding two digits (01, 02, 03, ...) at the end of the old household ID, in the order of which individual is found first (see example 2 in the Appendix Figure). One complication in this case could occur when the household members who move to different locations are tracked by different interviewer teams and the teams do not communicate with each other.
- The third option is to assign a new household ID to all households by using the old household ID and adding to that the roster number of the tracked respondents in that household (01, 02, 03.... up to however members were in baseline), as illustrated in example 3 in the Appendix Figure.

All of the options above make it easy for survey users to link the household across survey rounds by looking at the digits before the last two digits. The enumerators should then assign a person ID for each individual in the new roster. The person ID for each individual should be unique within the household and should follow a simple rule. For example, the household head gets person ID 1, the spouse gets 2, and so on. This new person ID is specific for the survey round and is not meant to be used to link individuals between survey rounds.

Creating unique IDs to identify individuals across rounds is the responsibility of the survey planners and should be done at the baseline. In the new household roster there should be a column for interviewer to put this cross-wave ID.

### **Completing a new household roster.**

Survey planners must decide on which information needs to be pre-loaded/pre-printed into the new household roster. Some options are as follows:

- The interviewer could start with a blank household roster and begin by writing down (or loading) the information about the tracked person who was found. Other household members are then added to the roster. An alternative would be to list everyone who was in the same roster as the tracked respondent, have a column to identify whether these individuals are in the households, and also have blank lines to add new household members previously not listed in the roster. In CAPI, it is possible to pre-load a list of all individuals who have ever been in the same households as the tracked individuals. The interviewer would then be able to select which of these individuals should be included in the new household roster. Survey planners must decide which basic characteristics will be forward-fed into the new household roster (such as age, sex, birth date) and what needs to be done when there is conflict between current responses and pre-loaded information.
- Interviewer must identify who among the new household members are the interview targets, as defined in Section 4. The rule to assign interview targets must already have been established by the survey planners. For example, the rule may state that only the tracking targets are supposed to be interviewed. Or, the interview targets could include a respondent's spouse and children, even if they have never been part of the survey before. In MHSS2 (as in IFLS) for example, only migrants, their spouses and their direct descendants are interviewed individually.

Suppose a tracking target moves to the household of his/her uncle. Should the uncle and his family all be interviewed? If not, who is to be interviewed for the household level information? Again, the rules in assigning this should be set by the survey planners and made clear to the interviewers.

## **10. Ethical concerns**

Longitudinal household surveys with tracking may be subject to some additional ethical requirements. The ethical implications of tracking activities, which include collecting information about future moves and following respondents who have moved, should be carefully considered.

Different surveys could have different options with regards to how to approach this issue, similar to the options faced by survey organizers on how to gain informed consent for a longitudinal survey without tracking. As an illustrative example, during the planning stage, the British Household Panel Survey faced the following options: 1) inform respondents that they were being recruited into a survey of indefinite

time period or fixed years, 2) only obtain consent of wave 1 without mentioning future waves, and 3) obtain consent for wave 1 while explaining the nature of longitudinal surveys (Lynn, 2006). The BHPS chose the third option and decided not to ask explicitly for consent for the following round, although the respondents would be informed that they would be visited again in the future.

Survey organizers who are planning to include tracking activities in the survey may follow a similar method by informing the respondents that they may be visited again in the future and if they move they may be followed to the new location.

## **11. Conclusion**

This paper emphasizes the importance of tracking respondents that have moved from the baseline households and communities when conducting longitudinal household surveys. Failure to follow movers entails scientific costs in terms of higher attrition rates, which stems from both observed and unobserved characteristics. While the financial costs of tracking can vary according to survey designs and settings, they can be minimized with careful planning and the use of innovative technology. This requirement implies that tracking calls for significant institutional capacity in regards to the organization of the survey. The implementation of successful tracking requires careful consideration and planning, as well as flexibility to new innovative approaches.

## References

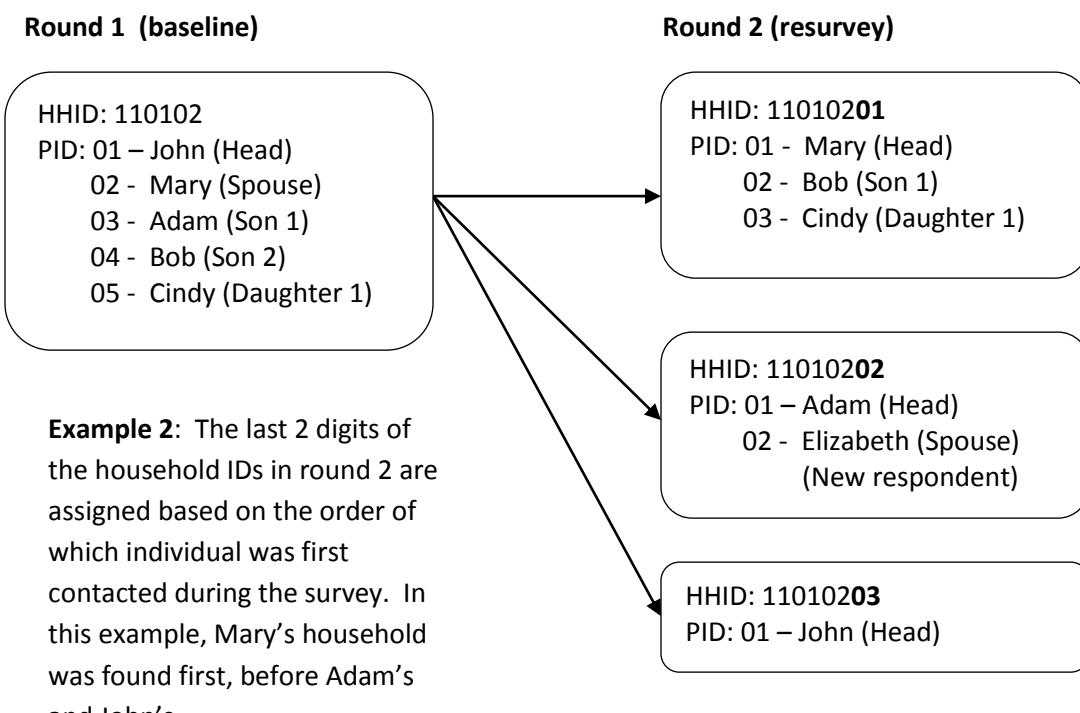
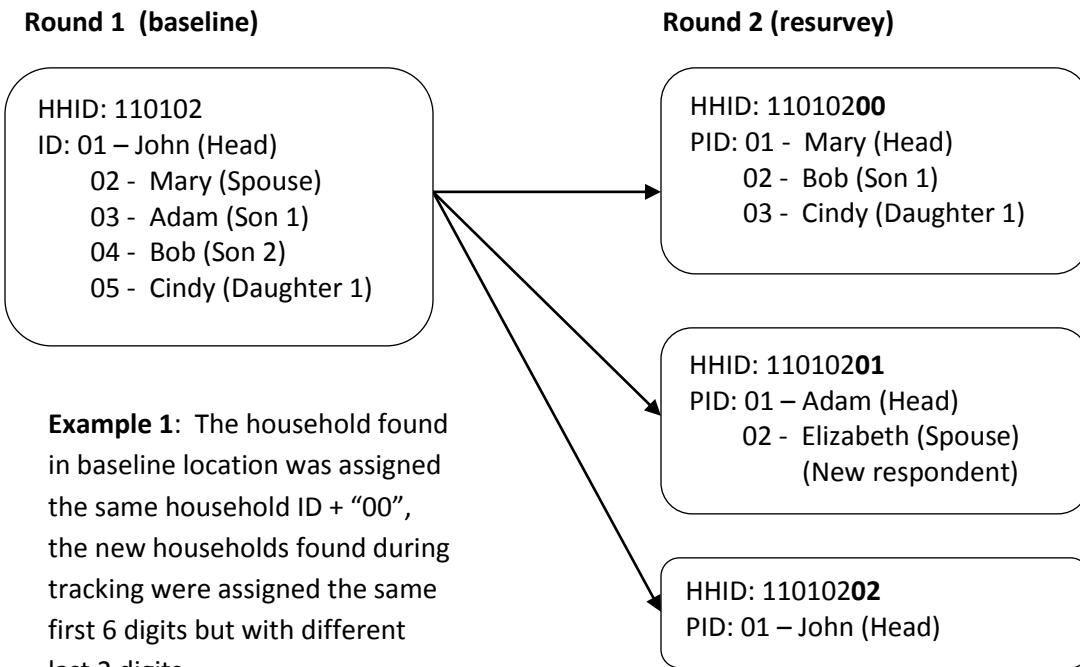
- Alderman, H., Behrman, J., Kohler, H.P., Maluccio, J.A. and Watkins, S. (2001). "Attrition in longitudinal household survey data: Some tests from three developing countries," *Demographic Research*, 5(4):79-124.
- Beckett, S., Gould W., Lillard, L., and Welch, F. (1988). "The Panel Study of Income Dynamics after fourteen years: An Evaluation," *Journal of Labor Economics* 6:472-92.
- Beegle, K., De Weerdt, J., Dercon, S. (2009). "Migration and Economic Mobility in Tanzania: Evidence from a Tracking Survey," *World Bank Policy Research Working Paper Series* 4798.
- Beegle, K., De Weerdt, J., Dercon, S. (2008). "Adult Mortality and Consumption Growth in the Age of HIV/AIDS," *Economic Development and Cultural Change*, 56: 299-326.
- Fitzgerald, J., Gottschalk, P. and Moffitt, R. (1998). "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics," *The Journal of Human Resources*, 33, 251-299.
- Frankenberg, E. and Thomas, D. (2000). "The Indonesia Family Life Survey (IFLS): Study Design and Results from Waves 1 and 2," RAND, Santa Monica, CA. DRU-2238/1-NIA/NICHD.
- Glewwe, P. and Jacoby, H. (2000). "Recommendations for collecting panel data", in M. Grosh and P. Glewwe (eds.), *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Study*. Washington D.C., World Bank.
- Hill, Z. (2004). "Reducing Attrition in Panel Studies in Developing Countries", *International Journal of Epidemiology*, 33, p.1-6
- Korinek, A., Mistiaen, J. A., Ravallion, M. (2007). "An Econometric Method of Correcting for Unit Nonresponse Bias in Surveys," *Journal of Econometrics*, Elsevier, vol. 136(1), pages 213-235, January.
- Laurie, H. and Lynn, P. (2009). "The Use of Respondents Incentives on Longitudinal Process", in Lynn, P. (ed) *Methodology of Longitudinal Surveys*. West Sussex: Wiley.
- Lynn, P. (2009). "Methods for Longitudinal Surveys", in Lynn, P. (ed) *Methodology of Longitudinal Surveys*. West Sussex: Wiley.
- Lynn, P, eds. (2006) "Quality Profile: British Household Panel Survey BHPS , Version 2.0: Waves 1 to 13: 1991-2003," Institute for Social and Economic Research.
- Rosenzweig, M. R. (2003). "Payoffs from Panels in Low-Income Countries: Economic Development and Economic Mobility," *American Economic Review* 93(2):112–17.
- Thomas, D., Frankenberg, E. & Smith, J. P. (2001). "Lost but Not Forgotten: Attrition and Follow-up in the Indonesia Family Life Survey," *Journal of Human Resources*, 36, 556-592.

Thomas, D., Witoelar, F., Frankenberg, E., Sikoko, B., Strauss, J., Sumantri, C., and Surastini, W. (2010)  
“Cutting the costs of attrition: Results from the Indonesia Family Life Survey,” *Journal of Development Economics*, doi:10.1016/j.deveco.2010.08.015

Uhrig, C. N. (2008). “The Nature and Causes of Attrition in the British Household Panel Study,” *ISER Working Paper* 2008-05

Witoelar, F. (2005). “Inter-household Allocations within Extended Family: Evidence from the Indonesia Family Life Survey,” *Yale University Economic Growth Center Discussion Paper* No.912.

**Appendix Figure. Household ID assignment: some examples**



**Round 1 (baseline)**

HHID: 110102  
PID: 01 – John (Head)  
02 - Mary (Spouse)  
03 - Adam (Son 1)  
04 - Bob (Son 2)  
05 - Cindy (Daughter 1)

**Round 2 (resurvey)**

HHID: 110102**02**  
PID: 01 - Mary (Head)  
02 - Bob (Son 1)  
03 - Cindy (Daughter 1)

HHID: 110102**03**  
PID: 01 – Adam (Head)  
02 - Elizabeth (Spouse)  
(New respondent)

HHID: 110102**01**  
PID: 01 – John (Head)

**Example 3:** The last 2 digits of the household IDs in round 2 are assigned based on the person ID of the individuals in the baseline survey, regardless of which household was found first.

Appendix Table. Longitudinal Surveys with Tracking in Developing Countries (sorted by the year tracking was first conducted in the survey)					
Study name and country	Baseline sample	Rounds	Initial study goals	Tracking strategies/features	References
Malaysian Family Life Survey; Malaysia	1,260+ ever married women below 50 and their households in Peninsular Malaysia	1976/77, 1988/89	to identify factors amenable to public policy influence that directly or indirectly affect fertility outcomes	started the follow-round in rural areas and worked towards urban areas (main migrant destinations), to reduce the need to return to study areas.	"The Second Malaysia Family Live Survey: Overview and Technical Report" (Haaga et al, 1993). Survey website: <a href="http://www.rand.org/labor/FLS/MFLS.html">http://www.rand.org/labor/FLS/MFLS.html</a>
China Health and Nutrition Survey; China	3,800+ households	1989, 1991, 1993, 1997, 2000, 2009; tracking in sample areas since 1993	to examine the effects of the health, nutrition, and family planning policies and programs affecting the health and nutritional status of its population	tracking limited to new households formed in the survey areas.	Survey website: <a href="http://www.cpc.unc.edu/projects/china/proj_desc/">http://www.cpc.unc.edu/projects/china/proj_desc/</a>
Nang Rong CBIRD Evaluation Project: Migrant Follow Up; Thailand		1984, 1994/95, 2000; migrant follow-up since 1995	to assess the impact of Community Based Integrated Rural Development project (CBIRD)	in 1995, 1800 migrants from 22 randomly selected villages (out of 51 study villages) were followed if they moved to major destination areas. Informants included at least 1 household member and village informants). Around 70% of target migrants were interviewed.	Survey website: <a href="http://www.cpc.unc.edu/projects/nangrong">http://www.cpc.unc.edu/projects/nangrong</a>
Matlab Health and Socio-economic Surveys (MHSS); Bangladesh	Main survey: 1996: 4300+ household Matlab, outmigrant survey: 550+ migrants who left the area between 1982 and 1996	1996, 2010; tracking in 1996 and 2010	studying the effect of socio-economic and behavioral factors on adult and elderly health status, and health care utilization	in 1996, tracked 550+ migrants who had left the area since 1982. MHSS2: tracked around 5,900 households originated from 2,700 primary households, plus up to 400 household overseas. Full time tracking manager hired to oversee tracking operation.	"The 1996 Matlab Health and Socioeconomic Survey: Overview and User's Guide" (Rahman, O. et al 1999) Email correspondence with Randall Kuhn (MHSS2 Principal Investigator).

Indonesia Family Life Surveys; Indonesia	7,000+ households in 13 provinces	1993, 1997, 2000, 2007; tracking since 1997	multi-thematic household survey	since 1997 tracking was conducted using multiple informants, centralized tracking management, and dedicated tracking teams. Tracking was conducted to outside enumeration areas within the survey provinces.	"IFLS2 Overview and User's Guide" (Frankenberg et al 2001), "IFLS3 Overview and User's Guide" (Strauss et al 2004), "IFLS4 Overview and User's Guide" (Strauss et al 2009). Survey website: <a href="http://www.rand.org/labor/FLS/IFLS.html">http://www.rand.org/labor/FLS/IFLS.html</a>
Vietnam Longitudinal Survey; Vietnam	1,800+ households	1995, 1996, 1997	to analyze the impact of the changing household economy on demographic phenomena in Vietnam		Survey website: <a href="http://csde.washington.edu/research/projects/hirschman/vietnam/INTROVLS.htm">http://csde.washington.edu/research/projects/hirschman/vietnam/INTROVLS.htm</a>
KwaZulu-Natal Income Dynamics Study; South Africa	1,350+ households in KwaZulu-Natal	1993, 1998, 2004; tracking in 2004	understanding income dynamics	in 2004, found 841 of the 1,350+ original dynasties. All eligible households were visited prior to main fieldwork.	The KwaZulu-Natal Income Dynamics Study (KIDS) 3rd wave: methods, first findings and an agenda for future research (May et al 2006)
Kagera Health and Development Survey; Tanzania	950+ households in Kagera	1991-1994, 2004; tracking in 2004	studying the impact of the health crisis linked to the HIV-AIDS epidemic in the area	pre-field tracking survey; multiple informants;	"User's Guide to the Kagera Health and Development Survey Datasets" (World Bank, 2004). "Kagera Health and Development Survey 2004 Basic Information Document" (Beegle, et al 2006)
Albania Living Standard Measurement Survey; Albania	1,700+ households, subsample of 2002 LSMS	2002, 2003, 2004; tracking in 2004		centralized tracking management, subset of international migrants to Greece were tracked	"Albania Panel Survey 2004" (INSTAT, 2005)
Mexican Family Life Survey; Mexico	8,400+ households, nationally representative	2002, 2006, 2009; tracking since 2006	multi-thematic household survey	tracking international migrants who moved to United States	"User's Guide: Mexican Family Life Survey 2005" (Rubalcava and Teruel, 2007); <a href="http://www.ennvihi-mxfls.org/en/mxfls.php">http://www.ennvihi-mxfls.org/en/mxfls.php</a>
Young Lives Project; Ethiopia, Peru, India, Vietnam	2 cohorts of children in each country; 2,000+ born in 2001/02, 1,000+ born in 1994/95	2000, 2007; tracking in 2007			"Survey Attrition and Attrition Bias in Young Lives", Young Lives Technical Note No. 5 (Outes-Leon and Dercon, 2008); <a href="http://www.younglives.org.uk/countries">http://www.younglives.org.uk/countries</a>

Cebu Longitudinal Health and Nutrition Survey; Philippines	cohorts of women giving birth in 1983/84 in Cebu: 3,000+ women in 2,600+ households	1983/84, 1991/92, 1994/95, 1988/89, 2002, 2005, 2007; "limited tracking" in 2007	to understand infant feeding decisions and consequences		"About the Cebu Longitudinal Health and Nutritional Survey" Survey website: <a href="http://www.cpc.unc.edu/projects/cebu/about.html">http://www.cpc.unc.edu/projects/cebu/about.html</a>
International Crop Research Institute for the Semi-Arid Tropics Village Level Study; India		1975-1985, 1989, 2001, 2007; tracking in 2007	to understand the farming systems in rural areas and identify socio-economic and institutional constraints faced by the farming community	tracking in 2007	
National Income Dynamics Study; South Africa	8,000+ households, nationally representative	2008, 2010	to understand the dynamics of income, consumption and expenditure of households		Survey website: <a href="http://www.nids.uct.ac.za/about-us.html">http://www.nids.uct.ac.za/about-us.html</a>
Kenya Life Panel Survey; Kenya	7,500 + young adults	Baseline in 1998, 6,800 children were tracked in 2008	to estimate the long-term impact of deworming project on educational, labor market, and health outcomes	tracking in 2008 followed children who moved out of the original project district in Busia	Survey website: <a href="http://cega.berkeley.edu/projects/kenya-life-panel-survey/">http://cega.berkeley.edu/projects/kenya-life-panel-survey/</a>