SWIFT Data Collection Guidelines version 2

June 9, 2015

Yoshida, N., R. Munoz, A. Skinner, C. Kyung-eun Lee, M. Brataj, W. Durbin and D. Sharma

Contents

1 Introduction

Limited data availability continues to thwart country efforts to monitor the World Bank Group's twin goals of "Ending Extreme Poverty" and "Promoting Shared Prosperity". **SWIFT (Survey of Well-being via Instant and Frequent Tracking)** is a powerful new survey package that can contribute to filling these persistent data gaps and to help project teams have a better understanding of who the poor are, where they live, and what constraints they face in escaping poverty.

SWIFT achieves this by combining the latest in Information and Communication Technology (ICT) and statistical methodology. SWIFT does not collect direct income or consumption data which can be both time-consuming and vulnerable to error without the right know-how and resources; instead, it collects poverty correlates, such as household size, ownership of assets or education levels, and then converts them to poverty statistics using estimation models. Collecting correlates is a far easier task that can be accomplished much more quickly than collecting income or consumption data.

In addition, SWIFT uses Computer-Assisted Personal Interview (CAPI) software to create and implement its questionnaires, allowing full and immediate accessibility from tablets and/or smart phones. Once enumerators have completed an interview, with the benefit of on-the-spot automated logic checks, data is uploaded to a data cloud, provided there is Internet access. Data analysts then download the data and convert them into poverty and distributional statistics. The kinds of findings produced depend on client demand. SWIFT can generate poverty and inequality statistics as well as poverty profiling and benefit incidence analysis.

Ultimately, these features allow SWIFT to estimate household income and expenditure data **in a cost-effective, timely, and user-friendly manner.** With the estimates, we can estimate poverty rates, inequality statistics, and income data. Furthermore, if a SWIFT survey is repeated a few years later, we can estimate growth of mean income over time and the shared prosperity index, which is a growth rate of mean income for the bottom 40 percent of population. In this way, SWIFT can help WBG staff align their projects and lending programs to the twin goals while also helping developing countries improve their poverty monitoring capabilities and design of pro-poor policies.

This document is intended primarily as a guide for those wishing to implement SWIFT without the guidance of the SWIFT team. As such, it contains a number of detailed technical sections on the theoretical background of the modeling and sampling processes. Section 2 however, is intended for users wishing to implement SWIFT with the help of the team. It contains a number of filter questions for teams considering SWIFT and subsequently lists the process and requirements under which the SWIFT team operates.

Sections 3-9 contain technical guidelines to carry out this type of survey. These guidelines are organized in six chapters, one for each part of the survey process: (i) modeling, (ii) questionnaire design, (iii) sampling, (iv) contracting and checklist (v) listing operation, (vi) staffing and training, (vii) fieldwork, and (viii) auditing. Each section has been divided into

sections, one for each step in the process. In most cases, steps are sequential, requiring the satisfactory completion of one step before moving on to the next.

Finally, there is an annex section providing examples of the various steps involved in a SWIFT project.

Since project teams will probably ask the assistance of respective National Statistical Organizations (NSOs), or hire a firm, to carry out many of the steps involved in the data collection process, the guidelines do not focus on how to carry out each step; the survey firm is better equipped to decide this (from hereon we will use the term *survey firm* to refer to the NSO or a private firm, indistinctively). Instead, the guidelines focus on how the project team can verify that steps are completed satisfactorily. However, to guide the survey in proper direction, a few steps need to be led by the project team, especially during the design stage. This includes holding meetings with stakeholders and the survey firm, where the team will establish the survey objectives and other requirements. For these steps, we provide guidelines on how the project team should proceed.

The SWIFT manual, which is now version 2, will continue to be updated as statistical methodologies and ICT technologies are improved. The core-team will keep examining the state-of-the art statistical methodology and ICT technologies and if they are found to be clearly beneficial, the team will update the SWIFT approach and the manual will be revised.

Project Team Requirements

2 How to Run SWIFT with the SWIFT team

SWIFT is a powerful survey instrument that can produce estimates of poverty and shared prosperity in a very timely and cost-effective manner. However, proper implementation requires careful preparation as mentioned. The SWIFT team can facilitate the preparation process. This section describes how to initiate the process and engage the SWIFT team.

The first and most important question is:

IS SWIFT FOR ME?

Project teams can assess whether a partnership with SWIFT would be valuable based on the following questions:

1. Does your project face concrete research questions relating to poverty? E.g. do you wish to
    a. test for differences in project outcomes between poor and non-poor beneficiaries, or bottom-40% and non?
    b. capture the increase in households' consumption or income level, or alternatively a shift in households' status from poor to non-poor, due to a project intervention?
    c. find out the incidence of poverty amongst your project's beneficiaries?
    d. select a target beneficiary group for a new project based on their poverty status?
2. Will you be administering, or are you willing to administer, a household survey to your project's beneficiaries in the near future?
3. Are you willing to include approximately a dozen or so questions which the SWIFT team will provide – straightforward questions with yes or no, multiple choice, or numerical response –to your survey?
4. Do you already have a sound sampling strategy for the survey, or are you willing to work with SWIFT to develop one?
5. *The SWIFT team can help you answer the following question if you are unsure*: Is there a relatively recent household expenditures/income survey in your project country for which the government has availed data?

2.1 Contact Us

The first step for anyone interested in SWIFT services is to contact the team by writing to any of the three below.

Nobuo Yoshida (nyoshida@worldbank.org)
Alexander Skinner (askinner@worldbank.org)
Spencer William Durbin (sdurbin@worldbank.org)

In this email it is very useful for the client to indicate, briefly, the nature of the project and how they hope to utilize SWIFT within that context. An attachment with a project brief is helpful as well. This will help the SWIFT team place the project into the various SWIFT sub-groups including SWIFT Finance, SWIFT Agriculture etc.

For more information, please visit our SPARK website at:

https://spark.worldbank.org/groups/poverty/projects/swift

2.2 Feasibility Test in terms of Data Requirements

As soon as the SWIFT team receives a request for the first consultation meeting, the team will investigate the availability of household survey data to assess the feasibility of conducting a SWIFT survey.

Data requirements for a SWIFT survey are fairly modest, but not negligible. An important issue is whether models created from the latest household survey data represent the current relationship of consumption or income. To ensure the stability of models, the SWIFT team would prefer to have at least two rounds of highly comparable LSMS (Living Standards Measurement Survey) or LSMS type multi-topic household surveys, no more than five years apart, with at least one of them being no more than three years old. Having two rounds of household survey allow the team to test whether the models developed from the latest household survey are stable over time.

In reality, less than half of available countries in the developing world have two comparable surveys over ten years. Consequently, many countries do not satisfy this condition. If this is the case, the following criteria are applied. First, if the latest survey was carried out inside two years or the survey is currently in the field, the SWIFT team produces models from the latest survey assuming that consumption patterns did not change much since the data were collected. Second, if the latest survey is too old, but there is a new survey in the field, the SWIFT team can create a questionnaire including variables that are likely to be in models that will be developed from the new survey. In this case, since the SWIFT team does not know the models yet, they need to include all potential candidates, which will increase the number of variables more than if the team knows the variables needed to conduct projections exactly.

However, our recent experience suggests that the need for including additional variables is minimal.

The SWIFT team will let the client know whether a SWIFT survey is feasible for monitoring or evaluating the impact of his or her project on poverty and shared prosperity and set up the first meeting with the client.

2.3 First Meeting

If the idea appears feasible, a SWIFT team member will meet with the client to determine the basic parameters that the SWIFT survey will take. This includes answering questions such as:

- What questionnaires are available?
- What kind of sampling frame does the client have in mind?
- What exactly does the project want to monitor with SWIFT data?
- What level of SWIFT service does the client have in mind?
- Does the client have any preference on the CAPI software?

2.4 Initial Analysis and Consultant Selection

A SWIFT team member will conduct an initial analysis of the available data that is intended for the survey. This is an integral part of the cost/time estimate. Amongst the factors likely to influence this are i) the types of variables available, ii) the extent of missing data, iii) the structure of the dataset and iv) quality of the data available. In particular, it is important that basic roster, asset and household characteristic data are in good shape as these are typically the types of variables used in SWIFT estimates.

At this stage, the SWIFT team will also look into its pool of experienced consultants to determine who is the best fit and most readily available for the job at hand. This consultant will then be presented to the client at the second meeting (Section 2.6)

2.5 Cost and Time Estimate

Once the basic facts are known to the SWIFT team, the team works out a rough timeline and cost estimate for the respective project. This estimate is then communicated to the potential client.

2.6 Second Meeting

If the potential client deems the estimate to be acceptable, the SWIFT team arranges for a second meeting to discuss any remaining issues, finalize the timeline and distribute tasks amongst the various team members and consultants. Once this is done, the survey can begin.

2.7 Hiring of the Survey Firm

The SWIFT team WILL NOT select the survey firm. That is the responsibility of the client. However, the client may choose to seek the SWIFT team's assistance in the hiring process. The SWIFT team will advise on important qualifications, such as CAPI experience, and provide supplementary input as needed based on past experience.

The client is free to make this choice without the SWIFT team as well, should they have preference for a specific firm for any reason.

For further details please refer to Section 3 of the manual.

2.8 Survey Implementation

Upon request, the SWIFT team will provide training for enumerators so that they can ask questions and learn how to handle the CAPI systems. The SWIFT team can also provide assistance in other areas such as sampling. Please refer to Sections 5 and 8 of the manual for details.

2.9 Data Management

Depending on earlier agreements and the type of CAPI software selected, the SWIFT team will provide the client with all necessary information to access the data being collected by surveyors. The SWIFT team will also be able to access the data from headquarters.

2.10 Delivering Results

The SWIFT team will generally provide an approximately 5-page summary of results including basic welfare statistics and details on the formulas used to produce these. It will also include a few paragraphs explaining the results in the context of the project in question. Depending on the earlier discussions, the type and detail level of this report can be adjusted to reflect the client's needs.

Technical Guidelines

3 Model Development

3.1 Basics and Assumptions

SWIFT collects only 10 to 20 questions on poverty correlates, projects household income or expenditure from them using a model, and estimates poverty and inequality statistics from the projected income or expenditure data. The poverty correlates usually include household size, household head's educational attainment, household head's employment status, ownership of consumer durables, housing conditions, etc. To do this accurately, model development is critical.

The model is developed assuming the relationship between household income or expenditure and poverty correlates is linear and also that there is an error in projection.[1] Equation (1) shows this relationship:

$$\ln y_h = x_h{}'\beta + u_h \quad (1)$$

where $\ln y_h$ refers to a natural logarithm of household income or expenditure of household h, x_h is a $(k \times 1)$ vector of poverty correlates of household h, β is a $(k \times 1)$ vector of coefficients of poverty correlates, k is a number of variables, and u_h is a projection error. In principle, SWIFT estimates the linear formula by regressing the natural logarithm of household income or expenditure on a set of poverty correlates in a household survey data that include both household income/expenditure and poverty correlates. The regression model becomes a formula, with which household expenditure or income will be projected into a dataset that has only poverty correlates. The latter dataset will be collected by a SWIFT survey.

The SWIFT modeling process includes multiple steps to improve the ability of the formula to project household income or expenditures by adjusting the coefficients (β) and estimating the distributions of both the coefficients and the projection errors.[2] No formula is perfect; so inclusion of the projection error is essential. Indeed, estimating the distribution of the projection error is key for estimating poverty rates and their standard errors.

[1] This does not mean SWIFT does not use a non-linear model, but it means that SWIFT's formula is linear in variables created in the dataset. Since some variables can be squares of other variables, SWIFT's formula can be non-linear. One of typical examples is that SWIFT uses household size and household size squared in a formula.

[2] The approach adopted by the SWIFT team is rather conservative in that the team did not adopt some approaches discussed at the frontier of research on modeling because the team thought evidence for these approaches is not yet strong enough. However, the team has been exploring such new techniques and may update the SWIFT modeling process once enough supportive evidence for these methodologies is provided.

3.2 Cross Validation

Since consumption patterns can differ significantly across areas and population groups, the SWIFT team makes efforts to create a model that is specific to the areas and population groups that a client is interested in. Such an adjustment is good to create the model tailored to the client, but can cause potentially large bias in poverty estimates because the sample used for creating a model declines by focusing on the specific group of population. "Over-fitting" is one of such problems. The over-fitting problem means that while a model can perform well within the sample developed for the model, it can perform badly outside the dataset. In a sense, the model over-fits the dataset used to develop it. To detect the problem the SWIFT team conducts a cross-validation analysis. The cross-validation approach separates data used for developing the model from those used for evaluating the model fitness.

More specifically, a household survey dataset is split randomly into 10 subsamples. Each of these subsamples is called a "fold." A consumption model is estimated from nine folds by running a stepwise Ordinary Least Square (OLS) regression.[3] The stepwise OLS regression means that a statistical package searches for an OLS regression model where all variables are statistically significant at a level set by an analyst. We use STATA and its stepwise selection model (described in detail in Annex 8). The nine folds used for developing a model are known as "Training Data".

After a model is selected, household expenditure or income data is projected using the model in the remaining fold, and a poverty rate and mean squared errors (MSEs) are estimated with the projected data. At the cross-validation stage, we project household expenditure or income data assuming the error term and regression coefficients follow normal distributions.

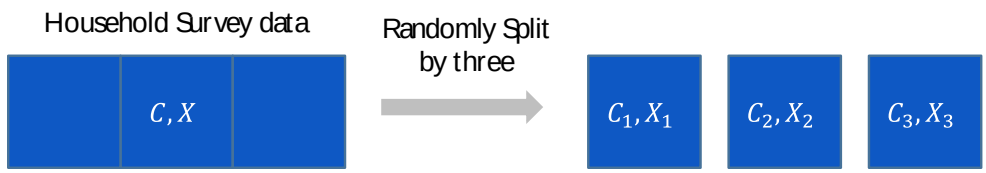
More specifically, suppose $\hat{\beta}$ is a vector of estimated coefficients and $\hat{\sigma}^2$ is an OLS estimator of error variance. We first draw a random value χ from a chi distribution with a degree of freedom, $(N-k)$, where N refers to the total sample size and k refers to the number of variables selected by the stepwise regression procedure, and calculate $\tilde{\sigma} = \hat{\sigma}(N-k)/\chi$. We then draw $\tilde{\beta}$ from a normal distribution of $(\hat{\beta}, \tilde{\sigma}(X'X)^{-1})$ where X is a $(N \times k)$ matrix of $(x_1, \dots, x_h, \dots, x_N)'$. Finally, we draw a simulated household expenditure or income for household h, $\widetilde{\ln y_h}$, from a normal distribution of $(X\hat{\beta}, \tilde{\sigma} I_{N*N})$ where I_{N*N} refers to an $(N \times N)$ identity matrix. This simulation process is repeated for all households, typically twenty times.[4] A poverty headcount rate is calculated by comparing the simulated household expenditure or income with a poverty line for each of the twenty simulation rounds. The average poverty rate of the simulations is used as a poverty estimate. MSE is calculated in testing data by taking the average of the sum of squared differences between y_h and $\hat{y}_h = x_h' * \hat{\beta}$.

[3] Or weighted least squares.
[4] This process can be done using STATA's command "mi impute regress", or STATA Corp LP (2013).

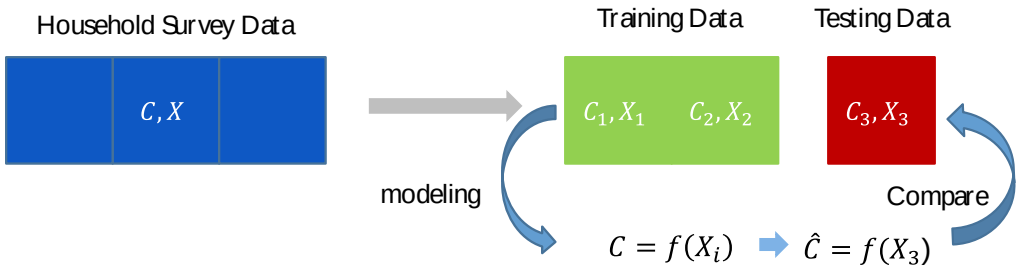
This analysis is repeated 10 times, each of which uses a different fold as testing data to test the performance in terms of mean squared errors and the absolute value of the difference between the projected and actual poverty rates. This test detects the over-fitting problem because all testing statistics are calculated from out-of-sample. SWIFT creates 10 folds. However, any number of folds can be chosen. Figure 1 shows an illustration of a three fold cross validation exercise.

Figure 1: Illustration of 3 Fold Cross-Validation

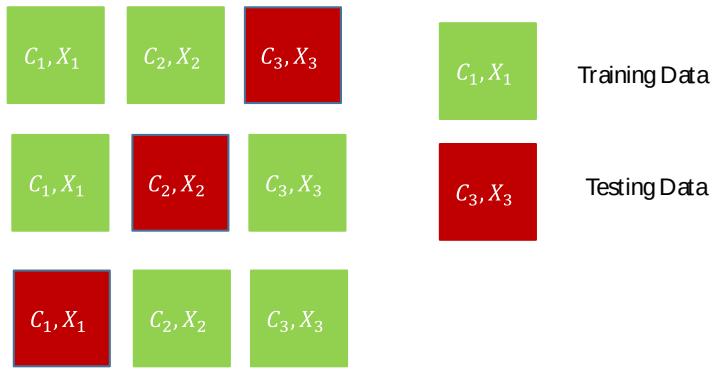
Step 1:Randomly split data into three folds (C refers to consumption; X refers to non-consumptiondata)



Step 2: Select two folds as training data, develop a model there, and test model performance in the testing data



Step 3: Repeat the above procedure three times by changing the testing data



This cross-validation exercise is conducted to determine the optimal threshold of the p-value for the stepwise regressions. For a specific p-value, the cross-validation exercise is done and produces the two testing statistics. The exercise is repeated for different levels of p-value, usually between 0.1 percent and 10 percent. The optimal p-value is the value where the absolute value of the difference between the actual and the projected poverty rates is minimized. The mean squared error is also examined to check whether the over-fitting problem occurs. If the mean squared error is minimized at a level of p that is smaller than the value where the absolute difference between the actual and the projected poverty rates is minimized, then the former value is chosen as the optimal number.

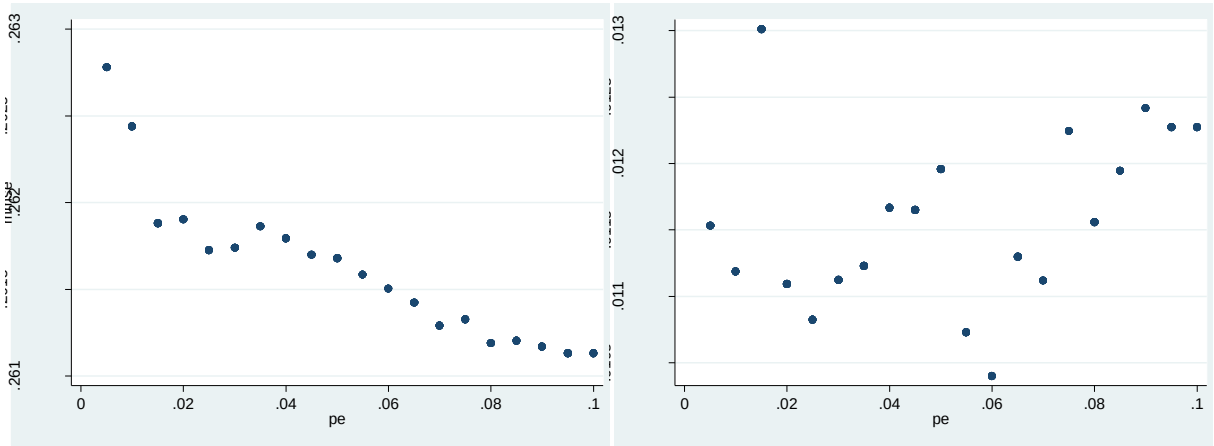
Figure 2 shows results of cross validation analysis using the Ghana Living Standard Survey (GLSS) 2012/13 data. The average MSE continues to decline as the threshold of the p-value for the stepwise regression increases. If MSEs are calculated in the same sample as where a model is developed, MSEs tend to decline as the p-value increases because the number of variables in a model tends to increase and the model fitness improves as the p-value increases. However, this is not always the case if we calculate MSEs out of sample because of the over-fitting problem. In the case of a cross-validation analysis for GLSS 2012/13 data, we did not see that, but we did see it in the other dataset. This suggests that there is no over-fitting problem in the modeling in GLSS 2012/13 for the range of p-values we investigated.

The average absolute values of the difference between actual and projected poverty rates show a different trend. Although the numbers fluctuate, it is clear that the difference starts increasing once the p-value reaches 6 percent. Below 6 percent, the value fluctuates, but it is never below the value at the p-value of 6 percent. Therefore, we choose 6 percent as the optimal threshold of the p-value for the stepwise regression procedure.

Figure 2: Typical Results of Cross Validation Analysis for Ghana 2012/13 data

Average MSE

Average absolute values of differences between actual poverty rates and projected poverty rates



Source: Results of cross validation analysis using GLSS 2012/13 data.

3.3 Finalizing the Model

After the optimal p-value is selected, a stepwise OLS regression procedure is carried out with a full sample of data to estimate a model. To ensure that the coefficients are stable, an OLS regression with the set of variables is carried out for all ten testing datasets to see whether the coefficients of the select variables do not change signs or are dropped due to collinearity. If some variables are dropped due to collinearity or some signs of the coefficients change, then these variables will be dropped from the final model. After dropping these variables, an OLS regression is carried out to estimate the coefficients and variance of the coefficients and error terms. In addition to the statistical tests, it is recommended to check whether the signs and values of all estimated coefficients make sense to those who know a country very well. If a sign of a variable is the opposite of an expert's intuition, this can be an indicator of multicollinearity and can be very unstable; therefore, it is strongly recommended to reconsider inclusion of such variables.

3.4 Simulation and Estimation of Poverty Rates

The final model is used to project household expenditure or income for all households 20 times following the procedure presented above. Poverty rates are estimated for each round of simulation and the average is taken as the estimate of the poverty rate. The variance of

the poverty estimate is calculated using the following formula (Rubin, 1987 and Schafer, 1999):

$$V(H^*) = \left(1 + \frac{1}{m}\right)\left[\left(\frac{1}{m-1}\right)\left(\sum_{l=1}^{m}(H^l - H^*)^2\right)\right] + \left[\frac{1}{m}\sum_{l=1}^{m} V(H^l)\right] \quad (2)$$

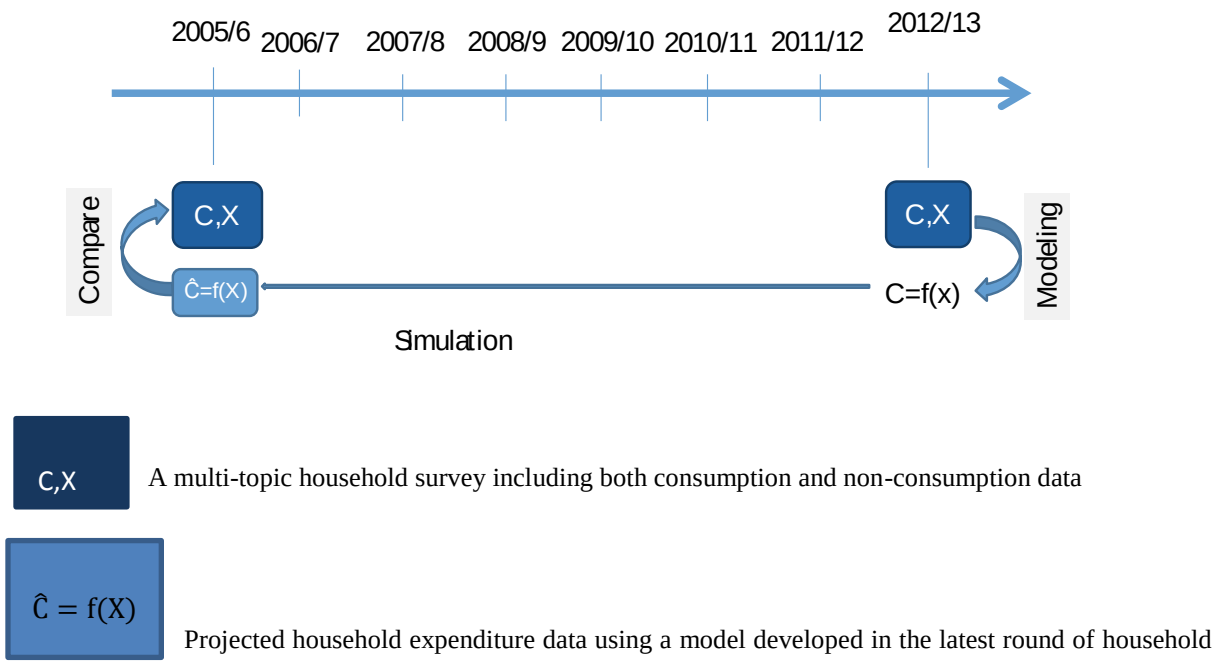
where m refers to the number of simulations, H^l refers to the poverty estimate in round l of the simulation, H^* refers to a mean of $\{H^l\}$ and the final estimate of the poverty headcount rate, m refers to the total number of simulations, and $V(H^l)$ is an estimate of the variance of the poverty estimate in round l of simulation. The first bracket presents the between simulation variance, while the second squared bracket presents the within simulation variance. Consequently, the variance of the final poverty estimate is a weighted average of the within and between simulation variances.

3.4 Robustness Tests

Finally, robustness tests are carried out. The first test is a test for model stability (or simply the "Stability Test"). The dataset used to develop the model will frequently have been collected two to three years before a SWIFT survey. Therefore, there is no guarantee that the model developed from a dataset in the past still represents a consumption pattern at the time the SWIFT survey is carried out. To test whether the final model developed after the cross-validation analysis is stable over time, a so-called "backward imputation" is conducted. The idea is to estimate a poverty rate using the final model of a previous round of household survey data.

Figure 2 illustrates the idea. A model (f(x)) is developed in the latest round of household survey (the 2012/13 data in this example). Then, household expenditure data are projected into the previous round of data (the 2005/6 data in this example) using the model. The simulation process is the same as that of the cross-validation analysis. Poverty rates are calculated with the actual consumption data and with the projected consumption data. We check whether the poverty estimate from the projected consumption data falls in the 95 percent confidence interval of the poverty estimate of the actual consumption data. If the model is stable over time, the projected poverty rate should be very close to that of the actual consumption data. Beyond the comparison of poverty rates, it is strongly recommended to examine summary statistics of variables for both the previous and the latest rounds of household survey data.

Figure 2: An illustration of backward imputation



Source: Authors' illustration.

The second robustness test examines whether assuming that the error term follows a normal distribution is valid (or simply the "Distribution Test"). All simulations in the above assume that the error terms follow a normal distribution. All distributions converge asymptotically to a normal distribution, but a SWIFT survey often focuses on a very small population. Therefore, assuming that the error term follows a normal distribution might be too strong an assumption. There is a possibility that the error term follows a distribution that is very different from a normal distribution. Also, the error term might include a cluster effect or heteroskedasticity. To see this, we adopt a simulation method developed by Elbers, Lanjouw and Lanjouw (2002, 2003), hereafter referred to as ELL.

ELL also assumes that the conditional expectation can be approximated by a linear model (1). However, it allows for a within cluster correlation in error terms:

$$u_h = \eta_c + \varepsilon_h,$$

where η and ε are independent of each other and uncorrelated with observables, x_h, and c refers to a cluster c. Furthermore, heteroskedasticity in the household specific error, ε_h, is also allowed. Taking these into account, the model in (1) is estimated by Feasible

Generalized Least Squares (FGLS). In order to estimate the FGLS model, the variance-covariance matrix, Σ is estimated as follows:

We first estimate the model (1) by OLS estimation. The residuals from this regression serve as estimates of overall disturbances, given by u_h. They are decomposed into uncorrelated household and location components:

$$\hat{u}_h = \hat{\eta}_c + e_h$$

The estimated location components, given by $\hat{\eta}_c$, are the within-cluster means of the overall residuals. The variance of the cluster effect, $\hat{\sigma}_\eta^2$, can be estimated by using a formula suggested by ELL. The household components of the error term, e_h are the overall residuals net of location components. To allow for heteroscedasticity in e_h, ELL proposes a logistic form:

$$\sigma^2(Z_h, \alpha, A^*) = \left[\frac{A^* e^{Z_h{}'\alpha}}{1 + e^{Z_h{}'\alpha}}\right]$$

where $A^* = (1.05) * \max\{e^2\}$. This form of heteroscedasticity model is restrictive in that the variance is bounded between zero and A^*, but is flexible in that a set of variables, Z_h, and their parameters, α can be estimated to fit the data. The optimal set of Z_h and the parameters are estimated from:

$$ln\left[\frac{e_h^2}{A^* - e_h^2}\right] = Z_h{}'\alpha + r_h$$

Letting $e(Z_h{}'\alpha) = B$ and using the delta method, the model implies a household specific variance estimator is:

$$\sigma_{\varepsilon,h}^2 = \left[\frac{AB}{1+B}\right] + \frac{1}{2}Var(r)\left[\frac{AB(1-B)}{(1+B)^3}\right]$$

Once this matrix has been calculated, the original model can be estimated by FGLS. The FGLS estimation produces a final set of first stage estimates for $\hat{\beta}_{FGLS}$, the coefficients from the main equation given by equation (1). The FGLS output also includes the associated variance-covariance matrix, given by $\hat{V}(\hat{\beta}_{FGLS})$, as well as parameters of the second round, $\hat{\alpha}$, $\hat{V}(\hat{\alpha})$, $\hat{\sigma}_\eta^2$, and $\hat{V}(\sigma_\eta^2)$.

Furthermore, ELL allows for distributions other than a normal distribution. After estimating the error distribution, coefficients, and distributions of coefficients, ELL simulates household expenditures by randomly drawing an error term and a set of regression

coefficients from the corresponding distributions estimated in the above. ELL usually repeats the simulations 100 times and estimates poverty rates for each round of the simulations. The final estimate of the poverty rate is the average of the poverty rates estimated from the 100 simulations. The simulation is done by PovMap 2 software – software the World Bank research department developed to conduct the above simulation procedure.[5]

To see the effect of introducing a more flexible distributional assumption, we run the above ELL simulation procedure with the set of variables in the final model, and check whether poverty rates projected by the ELL method fall into the 95 percent confidence interval of the poverty estimate of the actual consumption data in the latest round of household survey data as well as the previous round of household survey data. In addition, we compare the results with those estimated assuming the error term simply follows a normal distribution.

If the results of the model derived assuming the error term follows a normal distribution are not very different from those estimated from the ELL method, SWIFT picks the former model when estimating poverty rates in a SWIFT survey because handling models based on a normal distribution is much easier. In addition, the model based on the ELL method is sometimes unstable, and the model based on a normal distribution might outperform the other. As shown below, due to the instability, a model with a complex error structure may work well in projecting poverty rates in the latest round of a household survey, but it might not work as well in projecting poverty rates in the previous round of the survey.

Table 1 shows the results of the Afghanistan pilot where a model was created following the above procedure. A model was developed from the National Risk and Vulnerability Assessment (NRVA) 2011-12 survey data to project a poverty rate for NRVA 2013, which collected only non-consumption data. To correspond to the season when the NRVA 2013 survey was collected, a model was developed from a subsample of NRVA 2011/12 data. As a result, all poverty rates estimated from NRVA data below are not directly comparable to official poverty estimates.

For the distribution test, poverty rates were estimated under three different assumptions on the distribution of the error term – (i) a normal distribution; (ii) a flexible distribution without cluster effect or heteroskedasticity; (iii) a flexible distribution with cluster effect and heteroskedasticity. A flexible distribution implies that the error distribution is estimated using the semi-parametric estimation method available in PovMap 2, software developed by the World Bank Research Department (see more details in Lanjouw and Zhao, 2014). For the backward imputation, the poverty rates estimated in assumption (i) and (iii) are in the 95 percent confidence interval of the poverty rate estimated from the actual consumption data. The result under assumption (iii) slightly outperformed that of assumption (i). For the estimation of the poverty rate for the 2011/12 data, the poverty rates estimated in assumption (i) and (ii) are in the 95 percent confidence interval. The result under assumption (ii) slightly outperformed that of assumption (i). By looking at both results, a model developed under assumption (i) was selected to project a poverty rate using

[5] Software can be downloaded from http://go.worldbank.org/QG9L6V7P20.

NRVA 2013 data because the performance of poverty estimation under assumption (i) is consistently good under both estimations.

3.5 Summary

In sum, SWIFT modeling comprises the following steps:

1. Cross-validation to find an optimal level of p for the stepwise regression procedure

2. Definition of the final model using a stepwise regression with the optimal p-value as a threshold

3. Simulation of household expenditure or income data using the final model into the latest round of household survey and estimation of poverty rates following the multiple imputation method

4. Stability testing using "Backward Imputation" and model modification if necessary.

5. Distribution testing assuming a more flexible distributional assumption on the error term than a normal distribution using ELL's method and model modification if necessary.

6. Simulation of household expenditure or income data using the final (or modified) model into a dataset collected by a SWIFT survey, and estimation of poverty rates using the multiple imputation method.

Table 1. Comparison of results across different distributional assumptions in Afghanistan Pilot

(Model from the 2011/12 data)

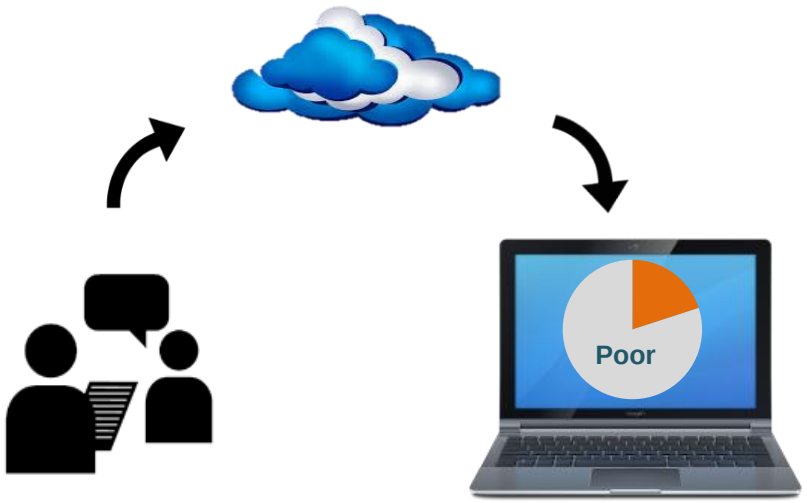
2007-08 (Estimation of contemporaneous poverty rates) – Backward imputation				
Actual		**Estimation**		
Poverty rate (%)	**95% CI**	**Poverty rate (%)**	**95% CI**	**Distribution**
37.4	[35.60 - 39.21]	37.53	[35.49 - 39.59]	Normal distribution
		41.41	[39.42 - 43.41]	Flexible distribution but without cluster or heteroskedasticity
		38.74	[35.67 - 41.81]	With cluster and heteroskedasticity
2011-12 (Estimation of contemporaneous poverty rates)				
Actual		**Estimation**		
Poverty rate (%)	**95% CI**	**Poverty rate**	**95% CI**	**Distribution**
37.47	[35.21 - 39.74]	36.43	[34.05 - 38.82]	Normal Distribution
		37.44	[35.01 - 39.85]	Semi-parametric estimation of error distribution
		34.51	[31.20 - 37.81]	With cluster and heteroskedasticity

Source: Afghanistan Pilot in Sherpa, Redaelli, Yoshida (2015).

4 Questionnaire development

4.1 CAPI platform

To reduce cost and increase efficiency and reliability, SWIFT utilizes the latest in Computer-Assisted Personal Interview (CAPI) technology. The process in itself is simple. Enumerators use either tablets or cellphones to collect answers to a questionnaire that has been created using CAPI software. Then, whenever they have access to 3G or wireless Internet, they upload the data to a data cloud. Supervisors can then download and check the data for mistakes or inconsistencies, sending the enumerators back for corrections or second interviews if necessary. Once the data are cleared, analysts from headquarters can download them and produce poverty rates, benefit incidence statistics and other calculations that may interest the respective project teams or National Statistics Offices. The diagram below illustrates this basic process:



SWIFT offers users the choice between two CAPI software products: SurveyCTO and World Bank CAPI. While both of these are highly effective data collection products, they have different characteristics that may make one or the other more suitable for a particular project.

SurveyCTO:

SurveyCTO is based on ODK (Open Data-Kit) and is the more flexible of the two. Its interface is highly user-friendly and it offers a plethora of options. With minimal programming and Excel based questionnaires, SurveyCTO can be learned in a short period of time.

SurveyCTO does not require a specific supervision structure, so questionnaires can be created quickly and with minimum background setup time. Questionnaires design is highly flexible so that it can be adapted to almost any specific wants and needs.

It is ideal for project monitoring and smaller surveys. If the premium is on speed, or if connectivity is limited, the SWIFT team recommends using SurveyCTO.

World Bank CAPI:

World Bank CAPI has a more defined structure. It offers rigid supervision and the questionnaires are created using its own (web) interface, rather than Excel. This means that, while there are fewer advanced features, data collection itself is more robust. Questionnaires appear in the headquarters' dataset only after supervisor approval. It takes longer to construct a World Bank CAPI questionnaire, since its interface requires individual clicking on every question. The software requires question-by-question input, so questions from previous surveys, or questions similar to one another cannot easily be copied.

SWIFT recommends this method for larger surveys with more long-term policy design implications, where there is sufficiently reliable connectivity for both enumerators and supervisors throughout the fieldwork, or when the need to oversee survey response quality is high.

Overall, both products offer a great variety of features in question design, sequencing, skip patters, GPS location technology, supervisory options etc. Surveys can also be constructed in a variety of languages. Both programs continue to be updated on a regular basis, correcting bugs and developing new features.

See Annex 7 for a more detailed comparison of the two methods.

4.2 Translating a SWIFT Model into a SWIFT Questionnaire

Once the modeling for the estimation of household consumption is complete, the next step in implementing SWIFT is to develop the questionnaire. The purpose of the questionnaire is simply to elicit the variables that make up the consumption model. The questionnaire may consist only of SWIFT questions – i.e. those corresponding to the variables selected in the model – or, more often, additional questions that the project team wants to include in order to measure project-specific results. Frequently, SWIFT questions are added to an already- planned project questionnaire in order to measure project results on groups of households with different expenditure levels.

The most important – and truly, only – rule in this step is to present the SWIFT questions exactly as they originally appeared in the survey upon which the model was built. The wording of the question and multiple choices, as well as the ordering of the choices, must be **exactly copied**. In rare circumstances, there may be exceptions to this rule – if, for instance, the survey must be conducted by SMS/text and the character limitations preclude the same wording – but this should be done very rarely, on a case-by-case basis, and with great care.

Example 1:

Suppose the SWIFT model includes the binary variable for whether the household's roof is made of tin. Suppose, furthermore, that this variable is constructed from the original survey question:

What is your roof made of?

1. *Tin*
2. *Concrete*
3. *Wood*
4. *Leaves*
5. *Other*

Note that, with categorical multiple choice questions, raw response data is often re-coded into multiple binary (dummy) variables in order to enable quantitative analysis. Associating "tin" with the value 1, "concrete" with value 2, "wood" with 3, etc. is essentially meaningless – i.e., wood is not three times the value of tin. Instead, constructing dummy variables for each type of roof material (e.g., "Is your roof made of tin? 1. Yes 0. No") allows for analysis.

When such a binary variable appears in the model, one must be careful to trace the original survey question, and avoid asking an altered question such as:

Is your roof made of tin?

1. *Yes*
0. *No*

There is a strong body of evidence suggesting that responses change when question wording changes, even if seemingly innocuously. Thus, we emphasize the importance of the rule to retain original questioning.

Note that certain variables included in the SWIFT model may require asking not one, but several questions in the survey instrument. Employment status, for example, is determined from a series of conditional questions as based on the definition of employment by the International Labor Organization (ILO). When a model includes employment status, the advantage is the ability to capture short-term, real-time change in well-being. Other household indicators, such as assets, are less elastic in the short-run. At the same time, adding the required ILO questions to the survey prolongs the duration of an enumerator's visit to each household. Thus to decide whether to include employment status in the survey or not, the project team should judge the trade-off between model precision and survey duration.

4.3 List of survey topics

The project team is the party mainly responsible for deciding on the survey topics. For example, a program team interested in providing a new treatment for malaria might make the following list:

- *Health status (self-reported symptoms or medical diagnoses)*

- *Access to health services and malaria treatment*
- *Use of mosquito nets, insecticides and other preventive interventions for malaria*
- *Water and sanitation*
- *Poverty*

This can also take the form of a set of policy questions that the team wants answered, for example:

- *Where do populations that most need treatment live?*
- *What barriers do these populations face in seeking treatment for malaria?*
- *Where are the main sources of mosquitos located?*
- *Where will the program be most effectively targeted to the poorer population?*

Remember that, ideally, a SWIFT survey is not intended to last more than 30 minutes per household interview. Testing the questionnaire will be essential to determine what can or cannot be done in practice.

4.4 List of survey indicators for each topic

The project team is also the main party responsible for developing a list of the indicators expected to be produced by the survey. The project team can resort to specialists and to the literature, if needed.

Below is a short example of how a list might look (although a real list could easily contain more indicators):

- *Percentage of children less than five years of age with reported fever in the last four weeks*
- *Percentage of individuals diagnosed with malaria (by a health professional) in the last 12 months, that received antimalarial drugs*
- *Average number of nights slept under an insecticide treated net, during the last seven nights*
- *Distribution of households by type of drinking water supply*
- *Poverty headcount index*

The project team might also want to compare some of these indicators across different subpopulations (for example, by district or by poverty status). A list of these subpopulations should also be made, in order to capture the necessary information to classify households in each group.

4.5 List of questions

Writing good questions is very difficult, and if the project team asks someone else to do this, we strongly recommend a close review of the work. Not doing so carries the risk of omission of essential information, ambiguity, misunderstanding, or bias, all of which may later foil indicators.

It is next to impossible to design good questions from scratch, and thus the project team should resort to the literature, and if needed, to specialists. To illustrate the importance of knowing the craft, consider that poverty researchers have been revising their questions and learning from their mistakes for decades. This happens in practically all other survey subjects, and shows that it is almost impossible to get it right the first time. In any subject, learning about previous experiences is paramount to avoid repeating mistakes.

Maintaining comparability with other surveys is another reason to review the literature. Project teams will often need to compare the results of the SWIFT survey with other surveys, and will have to use similar questions. It is also common to find that previous questions have deficiencies, so there is trade-off between maintaining comparability and improving the quality of the questions in your survey.

Some general good practices for writing questions include:

- **The reference period must be explicit**. For example, "During the last 12 months, did you receive antimalarial drugs?"
- **Questions should be asked one at a time**. Double-barreled questions should be avoided, as well as second questions hidden in the response options. For example, *"Did you receive antimalarial drugs in a hospital/health center, from a community health worker, from a friend, or did you have to buy it in a pharmacy?"* has four questions in one. It is not clear what will be recorded: whether the person received drugs at any of these places; or which place the person received the most drugs; or all the places where the person received drugs. It would be better to separate into more questions, for example, first ask *"During the last 12 months, did you receive antimalarial drugs?"* and then follow up with *"Where did you get the last antimalarial drugs you took?"* or *"During the last 12 months, did you receive drugs in a hospital/health center?"*
- Similarly, in cases where answer choices are read as part of the question, double-barreled answer choices should be avoided, such as "*No, I did not seek treatment because I had no time or money."* Instead, two choices should be offered: *"No, I did not seek treatment because I had no time"* and *"No, I did not seek treatment because I had no money"*, with the option to select more than one answer if appropriate.
- **Be straightforward and avoid double negatives**. For example, *"Did you fail to seek treatment for your fever?"* with answer options *"Yes, failed"* and *"No, did not fail"*, is not a good wording. A better wording would be *"Did you seek treatment for your fever?"* with answer options *"Yes"* and *"No"*.

- **Avoid biased or leading questions**, such as "Since it is good to have your fever treated, have you sought treatment for it?", or "You have sought treatment for your fever, haven't you?"
- **Avoid ambiguous questions** that different respondents might interpret in different ways. For example, consider the question *"Did you visit a health center in the last four weeks?"* Surely those visiting a health center to receive treatment will answer *"Yes"*, but it is not clear what those going to the health center for other reasons (ask for an appointment, accompany someone else, etc.) will respond.
- **Avoid technical language** that respondents might not understand.
- Questions should be culturally sensitive and appropriate.

It is also important to indicate which household member is expected to answer each question.

4.6 Questions numbered, coded, and organized into modules

This product consists of an Excel file, with one worksheet per questionnaire module, each worksheet containing a list of questions, numbered sequentially. The project team should review the following:

- Are the modules arranged logically (for example, by topic)?
- Is the title of each module appropriate?
- Is the respondent specified at the beginning of each module?
- Is the order of the modules appropriate? Tips:
    - If there are several respondents, modules for the same respondent should be together.
    - Start with the list of household members.
    - Try to leave sensitive questions for the end.
- Are the questions the ones agreed on in the previous product?
- For all questions with numerical answers: have the units been defined?
- For all questions with non-numerical answers: have the answer options been appropriately defined? Tips:
    - Many answers will consist of simply two answer options: *"Yes"* or *"No"*.
    - Other answers will require more options. Avoid splitting into options that you will not require at the data analysis stage, unless you are following the answer options spelled out in a previous survey.
    - Avoid splitting into options that the respondent will not be able to distinguish. For example, consider the question *"Where did you get the antimalarial drugs?"* and the following answer options:

        Hospital
        Health center
        Community health worker of MOH
        Community health worker of NGO
        Community health worker of WB Program
        Pharmacy

Friend/Family
Other

In the case that many respondents could not distinguish between the different types of community health workers, collapsing the three options into one is warranted.

- Check that each answer option has a unique code.
- Add the option *"Other"* only if needed, or if you don't know what answer to expect.
- Avoid putting *"Doesn't know/respond"* in the list of options. It may invite interviewers to use this option.
- Some questions may require multiple answer options. For example, the question *"What are all the treatments that you received for your fever?"* would have a list of treatment options from which the respondent could choose several. Since multiple answer options will become a set of dummy variables during data analysis, it is often worth considering reducing multiple options to only one option, by rephrasing the question as *"What was the <u>last</u> treatment that you received for your fever?"* or *"What was the <u>main</u> treatment that you received for your fever?"*
- Unless they are absolutely necessary, avoid open-end answers. These will be difficult to use during data analysis.

- Is the order of the questions logical, and are skips correctly defined? Tips:
  - Questions that only apply for certain individuals (for example, asking about seeking treatment for fever only applies to those who had fever) should come after a question meant to identify these individuals (first ask if the person had fever). If the person did not have fever, a skip should jump all the questions applying for those who had fever.
  - Questions should be numbered from 1 to *n*.
  - Skip instructions should be written next to the answer option triggering the skip, and should indicate the question number where to skip (or *"Skip to next section"*). If tablets are used, they should skip directly to the appropriate question.

- Some sections require repeating a set of questions, for a list of persons, businesses, food items, etc. This should be clearly indicated in the Excel file.

This will be a CAPI questionnaire, so the visual format of the Excel file is not important, as long as the project team is able to review it easily. Annex 2 shows a sample printout of a CAPI questionnaire.

If the questionnaire requires translation, then the texts – as exactly translated in other languages – should be included in the Excel file.

4.7 Automatic checks

Automatic checks in the CAPI questionnaire are one of the best tools to control quality, and thus the project team should review the list of checks. One type of checks consists in

sending warning messages for out-of-range responses, when, for example, the interviewer enters a number smaller than 2014 for the interview year. The survey firm should provide a list with the minimum and maximum ranges for all numerical responses. These are often included in the CAPI questionnaire printout, as can be seen in Annex 2.

Another type of checks consists in verifying the consistency between two or more responses, for example, checking that the age difference between mother and son is realistic. It is important that the project team contributes to this list as much as possible. Annex 3 shows a list of consistency checks for a sample questionnaire.

4.8 Alpha version of the questionnaire

Alpha is the first fully functional version of the CAPI questionnaire. After the Alpha version has been thoroughly tested by the software programmers, the project team should ask for demonstrations, and test the questionnaire themselves. One option is to ask the survey firm to test the Alpha version, in sessions accompanied by a member of the project team. Another option is for the project team to carry out independent tests.

The project team should review that the Alpha version of the questionnaire meets the following criteria:

- All the questions in section 4.6 have been included in the appropriate order.
- The wording of all questions and answer choices is correct.
- All skips are working.
- The automatic range and inconsistency checks are working.
- Questionnaire can be uploaded and downloaded with no issues.
- If multiple languages are being used, all of them work and display properly.
- Formatting (field-list, radio buttons, etc.) is correct.
- Mandatory questions cannot be ignored.
- The program does not crash or freeze.

Each problem found by the project team should be reported indicating clearly where the problem is, and how to replicate it on screen. Also, screenshots of the problems should be included.

4.9 Beta version of the questionnaire

The programmers should produce a Beta version of the questionnaire, which solves all the problems detected during the Alpha version tests. The survey firm and the project team should thoroughly test the Beta version in the field. It is here that the questions designed in section 4.6 will be truly tested with the population of interest.

In addition to the criteria used to test the Alpha version, the project team should have the following in mind when testing in the field:

- Has the questionnaire been tested in different situations (for example, different regions, different socioeconomic and education levels, etc.)?

- If there are skips, have all questions been tested?
- Are questions readily understood by respondents?
- Are there questions with little or no value that could be deleted?
- Does the questionnaire capture all the necessary or relevant information?

The project team should ask the survey firm to test the Beta version iteratively, that is, solving problems on a daily basis, so that each day an improved version of the software is used. These daily fixes include dropping and adding questions, changing the wording, changing answer choices, fixing skips and adding checks, among other things.

Finally, the survey firm should measure interview times per section. This is useful to shorten the questionnaire if needed.

4.10 Final version of the questionnaire

The Beta version tests should be done iteratively until all problems in the questionnaire have been fixed. This produces the final version of the questionnaire, which will be used in the field and during training.

The project team should ask for updated versions of the products of sections **Error! Reference source not found.** and **Error! Reference source not found.**, with the final lists of questions, answer codes, skips, checks, etc.

5 Sampling

Selecting a proper sample of households is essential to produce unbiased and precise survey indicators. The chief requirement is to obtain appropriate *sampling frame*, that is, an updated list of all households in the population of interest, from where the sample can be drawn. Without an updated sampling frame, some households in the population of interest will have a null probability of being selected, which biases the survey indicators. Conversely, an outdated sampling frame may contain households, which have moved away and are no longer appropriate to be selected.

Unless a census has been carried out less than 3 months before the survey, an updated list of all households in the population will not be available beforehand. The survey firm will have to carry out sampling in two stages. In the first stage, it will have to obtain a list of all *census enumeration areas* that make up the population (also called *clusters*, which are groups of roughly 100 neighboring households), and select a random sample of clusters with probability proportional to size (PPS); the size of a cluster is the number of households in it as per the latest census. This list of all the clusters in the population is called *sampling frame of primary sampling units* (PSU).

In the second stage, the survey firm will have to produce the *sampling frame of secondary sampling units* (SSU), which consists in an updated list of all households in each selected cluster. For this, the firm will have to carry out a *listing operation* in each selected cluster.

From the updated lists, a fixed number of households must be selected randomly in each cluster, regardless of the size of the cluster.

This section provides guidelines for the review of all products associated with sampling, except for the listing operation, a survey in its own right, which is covered in Section 7. This section covers six products associated with sampling, which are delivered in the following order: (i) sampling design, (ii) sampling frame of PSUs, (iii) sample of selected PSUs, (iv) sampling frame of SSUs, (v) sample of selected SSUs, and (vi) sampling report. The first three must be delivered before the listing operation. The last three can only be delivered after the listing operation.

5.1 Sampling design

The project team should closely oversee the sampling design, which consists in determining the sampling precision of the survey. Sampling precision is also the main survey cost factor, so it is important for the team to understand the sampling strategies available.

5.1.1 Stratification

The project team will have to lay out the stratification requirements for the sample, if any. Sample stratification means separating the population into subgroups, or strata, and then drawing a sample separately in each stratum. There are several types of stratification, with different objectives:

- **Proportional allocation.** The sample size per stratum is proportional to the population size per stratum. It normally gives good sampling precision for the total sample, and for the bigger strata, but gives bad precision for the smaller strata.
- **Equal allocation.** The sample size is the same in all strata. Equal allocation helps achieve similar precision across all strata, regardless of their sizes. The trade-off is generally a lower precision for the total sample, compared with the proportional allocation.
- **Optimal allocation** (also known as Neyman's allocation). If the project team can obtain information on the variance of indicator variables and/or the interview unit cost, by stratum, then it is possible to make an optimal allocation, which minimizes the sampling error of the total sample and/or the survey cost. If there is no specific information by stratum, then proportional allocation will minimize the sampling error of the total sample.
- **Markwardt's allocation.** Simply calculate the average sample sizes obtained from proportional and equal allocations. This will give you a compromise between both types of stratification.

If the project team deems that stratification is necessary, then they should specify which type.

5.1.2 Sample design report

The sampling design should be presented in a report, to be reviewed by the project team. The review is mainly quantitative; a good sample design report should contain the following information:

- If stratification was asked for, a table indicating the sample allocation, by stratum, and stratification weights. Table 1 shows two examples (proportional and equal allocation). With proportional allocation, sampling weights are the same in all strata, and can be omitted. With equal allocation, sampling weights vary by stratum, and should be presented in the report.

Table 1 Example of sample stratification

Stratum	Number of households in the population	Stratification with proportional allocation		Stratification with equal allocation	
		Sample size	Sampling weight	Sample size	Sampling weight
North region	230,000	230	10,000	260	8,846.2
East region	90,000	90	10,000	260	3,461.5
South region	550,000	550	10,000	260	21,153.9
West region	170,000	170	10,000	260	6,538.5
Total	1,040,000	1,040	10,000	1,040	10,000.0

- An explanation of how the sampling frame of clusters will be obtained, and what information it will contain, such as cluster size, and stratum variables.
- A mention that clusters will be selected with PPS, and that a fixed number of households will be selected in each cluster.

The report should also include the following information for each stratum:

- The stratum sample size, and the number of selected households per cluster (m).
- Standard error estimates (the standard error is a measure of the expected sampling error) for the indicators asked for by the project team. The standard error estimates should show the following information:
    - The number of expected observations for the variable used to generate the indicator (n)
    - Variance of the variable (σ^2)
    - Intracluster correlation coefficient of the variable (ρ)
    - Predicted standard error (e).

Also the standard error for the total sample should be reported.

The project team should check the predicted standard error in each stratum, using the following formula:

$$e = \sqrt{\frac{\sigma^2}{n}\left(1 + \rho(m-1)\right)}$$

The values of $\boldsymbol{n}$ and $\boldsymbol{m}$ should reflect the number of expected observations used to generate the indicator, which is not necessarily the total number of households in the sample. For example, the indicator "Percentage of individuals receiving antimalarial drugs, among those diagnosed with malaria (by a health professional) in the last 12 months" should be calculated over individuals with malaria only. In the formula above, n and m should be estimations of the number of individuals with malaria in the total sample and per cluster, respectively.

The result of the formula does not have to be exactly equal to the standard error in the report, but should be similar.

To check the standard error for the total sample, the following formula can be used:

$$e = \sqrt{\left(\frac{N_1}{N}e_1\right)^2 + \left(\frac{N_2}{N}e_2\right)^2 + \cdots + \left(\frac{N_s}{N}e_s\right)^2}$$

where N_i and e_1 are the population size and standard error in stratum i, respectively, and N is the total population size.

The number of selected households per cluster should be between 10 and 20. It is a good idea to ask for two versions of the sampling design, one with 10 and the other with 20 households per cluster, but both with the same resulting standard errors. Choose the one with lower cost.

5.1.3 Use of replacement households

Some surveys include an extra sample of households to use as replacements in case of non-response. Project teams should try to avoid the use of replacements, because they create perverse incentives: interviewers are supposed to replace households that really cannot, or will not, respond, but they will tend to misemploy the replacement sample if that makes the job easier for them, actually increasing the non-response rate and indicator bias.

Not using replacements has the disadvantage that the resulting sample size is uncertain. It helps to have some information on the expected non-response rate, to inflate the sample size correspondingly. For example, if a sample size of 1,000 households is desired, and the expected non-response rate is 10 percent, then the survey should be planned for 1,111 households. Of course, the resulting sample size may be smaller or larger than 1,000.

Not using replacements, and inflating the sample size, is the preferred option. Even if the resulting sample size turned out to be smaller than expected, a loss in sampling precision is generally less undesirable than an increase in bias.

The survey firm may have concerns on the cost associated to inflating the sample size. The cost does increase if the inflation is done by increasing the number of selected clusters. However, the cost does not increase if the inflation is done by increasing the number of households selected in each cluster.

Regardless of whether replacements are used or not, it is fundamental that the survey firm keeps a record of households that did not respond. This record will be needed for the calculation of non-response weights (see Section 5.6.3).

5.2 Sampling frame of PSUs

The project team should review the sampling frame of PSUs, which is the list of all clusters in your population, in an appropriate electronic format. The sample of clusters for the survey will be drawn from this list.

These sampling frames generally come from the census, and are often considered confidential. Even if the files cannot be distributed, the project team should ask to make a review of the how the sampling frame is organized, in the premises of the census agency if needed.

Table 2 shows an example of a properly organized sampling frame. The following criteria should be used to review it:

- Does the sampling frame consist of a list of all the clusters that make up the population of interest?
- Does each row represent a cluster? In the example, each cluster is a census enumeration area.
- Does each cluster have a unique identifier? In the example, each cluster is identified uniquely by the combination of District code and Enumeration area.
- If stratification is required, does the list include the variable(s) indicating to which stratum each cluster belongs to? In the example, variable Stratum would allow to stratify by region before drawing the sample of PSUs.
- Does the list have a variable indicating the size of each cluster, given by the number of households as per the latest census, or from another reliable source?
- Is the sum of variable Size consistent with the total number of households in the sample design report? In the example in Table 1, the total number of households is 1,040,000, and the sum of Size should be very close to this number. The same should be checked per stratum.

Table 2 Example of sampling frame of PSUs

Stratum	District code	Enumeration area	Size (number of households as per last census)
North region	101	1	127
North region	101	2	112
…	…	…	…
East region	201	1	98
East region	201	2	85
…	…	…	…

The project team should also ensure that the census agency will collaborate in providing the information to locate the selected clusters in the field.

5.3 Sample of PSUs

The sample of PSUs should be presented in a data file format (Excel, Stata, or similar), with the same structure as the sampling frame (Table 2). The only difference is that the list will contain the selected clusters only.

The project team should review the file with the same criteria used to review the sampling frame, except that the number of selected clusters in each stratum should be consistent with the sample sizes set forth in the sampling design.

5.4 Sampling frame of SSUs

The project team should review the sampling frame of SSUs, which is an updated list of all households in the selected clusters, in an appropriate electronic format. The sample of households for the survey will be drawn from this list. This sampling frame is the main product of the listing operation, and should be reviewed following the guidelines in Sections 7.6 and 7.7.

5.5 Sample of SSUs

The sample of SSUs is the final list of households that will be surveyed. The project team should verify that the number of clusters, and the number households per clusters, is consistent with the sampling design. Also, the list should have the same information as the sampling frame of SSUs (household head name, address, etc.)

It is fundamental that each household in the sample is assigned a unique identifier. Normally, the household identifier is the combination of the cluster's unique identifier and the household number from the listing operation. This household identifier will be used to identify each interview during fieldwork and in the final dataset. The project team should

ensure that the unique household identifier is included in the final sample file and in the questionnaire.

5.6 Sampling weights

The project team should also review the calculation of sampling weights to be used during the analysis. Three types of sampling weights should be taken into account: stratification weights, weights to correct for PSU size, and non-response weights.

5.6.1 Stratification weights

Depending on the type of stratification, sampling weights may or may not be needed during data analysis. Stratification with proportional allocation produces a self-weighted sample, so weights are not needed. Other types of stratification require that each household in the sample be assigned a weight, depending on the stratum it belongs to. The weight for each stratum is equal to the stratum population size divided by the stratum sample size.

5.6.2 Weights to correct for PSU size

After the listing operation (see Section 7) the survey firm will have updated the number of households in each PSU. In PSUs where the updated number of households is different than the number of households as per the last census, weights will need to be calibrated. The calibration consists in adding two columns to the sample of selected PSUs in Table 2. The first column contains the updated size of each selected cluster, that is, the number of households counted in the listing operation (excluding non-dwelling structures). The second column contains the weight correction factor, calculated as the updated size divided by the size as per the latest census (see Table 3).

Table 3 Example of sampling weights corrected for PSU size

Stratum	District code	Enumeration area	Size (number of households as per last census)	Updated size (number of households listed)	Weight correction factor
North region	101	1	127	135	1.063
North region	101	2	112	124	1.107
…	…	…	…	…	…
East region	201	1	98	101	1.031
East region	201	2	85	82	0.965
…	…	…	…	…	…

5.6.3 Non-response weights

If replacement households are not used (as recommended above) then the project team should correct for non-response. This consists in inflating the weights so that results expand to the entire population, in spite of some households not responding to the survey.

The first step is to calculate the response rate (the number of households who responded to the survey divided by the total number of households in the sample) for different types of households, for example by strata or by PSU. These response rates should be calculated using the stratification weights and the weights to correct for PSU size.

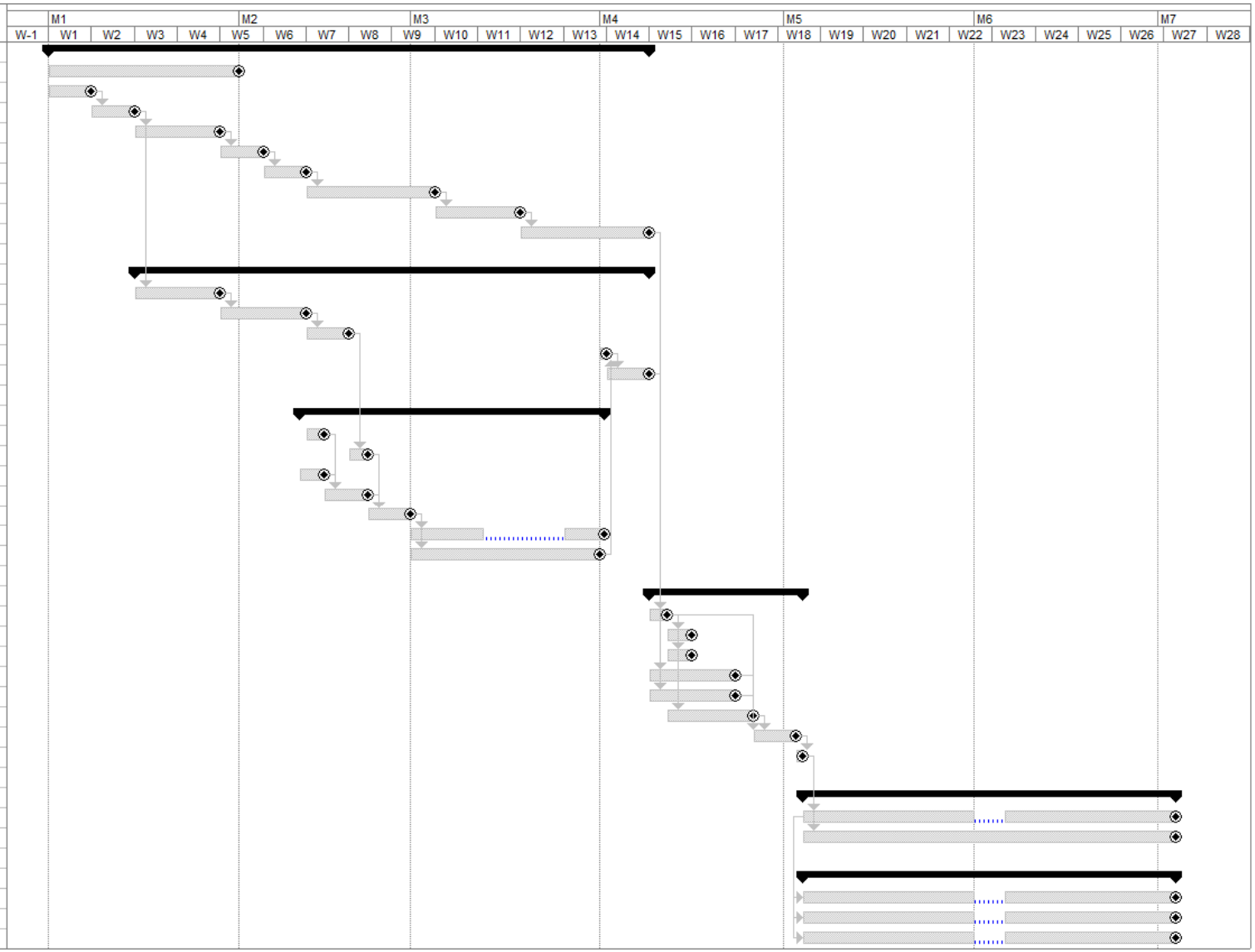
The second step is to calibrate the weights by a factor equal to the inverse of the response rate.

6 Contracting/Agreement

As we will show throughout this document, the project team will require several deliverables from the survey firm, and it is important to incorporate them into the survey firm's Terms of Reference (TOR). This section contains a checklist of all these deliverables, although project teams may choose to ask the survey firm for some of these products only, and develop the rest themselves.

The timeline on the right shows suggested landmarks for each deliverable (day 1 is the day you sign the contract with the survey firm). The bars represent tasks leading to each deliverable (task durations are referential). Some tasks may be conducted in parallel, but most must be conducted in a logical order, which is reflected in the timeline. You should cross out the items in the list in this order.

#	Task Name
1	Questionnaire development
1.1	CAPI platform
1.2	List of survey topics
1.3	List of survey indicators for each topic
1.4	List of questions
1.5	Questions numbered, coded, and organized into modules
1.6	Automatic checks
1.7	Alpha version of the questionnaire
1.8	Beta version of the questionnaire
1.9	Final version of the questionnaire
2	Sampling
2.1	Sampling design
2.2	Sampling frame of PSUs
2.3	Sample of PSUs
2.4	Sampling frame of SSUs
2.5	Sample of SSUs
3	Listing Operation
3.1	Listing form
3.2	Listing operation calendar and staffing plan
3.3	Listing operation supervision plan
3.4	Listing operation manual
3.5	Delivery of listing operation training
3.6	Listing operation weekly reports and datasets
3.7	Listing operation final report and dataset
4	Staffing and Training
4.1	Calendar and staffing plan
4.2	Sample coverage plan
4.3	Supervision plan
4.4	Interviewer and supervisor manuals
4.5	Training materials and evaluations
4.6	List of interviewer and supervisor candidates
4.7	Delivery of training
4.8	Selection of field staff
5	Fieldwork
5.1	Weekly reports and datasets
5.2	Final report and dataset
6	External field supervision (suggested)
6.1	Telephone check-up calls
6.2	xx
6.3	xx



The project team should also try to incorporate proper incentives in the TOR, to counteract the fact that survey firms tend to receive an important portion of the total payment before having delivered any data, which puts them in a good bargaining position to lower the quality of fieldwork. These payments are for preparation activities, such as the development of the questionnaire and interviewer training. Once these payments are made, the failure cost falls on the project team. Consequently, obtaining adequate solutions from the firm can be difficult when problems arise in the field, as it often happens. Some recommendations are the following:

- Try to balance payments in favor of the project team, by paying a large portion of the total payment after delivery of the data.
- Do an external quality control during fieldwork.
- Make one payment proportional to the coverage achieved.
- Ask for the delivery of partial datasets weekly during fieldwork.

To assess if the economic proposal presented by the firm is reasonable, the project team can use a budget template included in this guide (see Annex 1). The template includes the budget items required for a survey. It also allows you to calculate fieldworker transportation, salaries and per diem costs, based on survey parameters.

7 Listing Operation

The main objective of the listing operation is to produce the sampling frame of SSUs, which is an updated list of all households in each selected cluster. Another objective is to produce instructions on how to locate each cluster and each household. Listing is performed in the field by the survey firm.

The listing operation is a survey in its own right, and several products need to be reviewed by the project team if it is to be done correctly. A poor listing operation will translate into a poor sample of household, ultimately causing bias.

The Demographic and Health Surveys (DHS) Sampling and Household Listing Manual (http://dhsprogram.com/pubs/pdf/DHSM4/DHS6_Sampling_Manual_Sept2012_DHSM4.pdf) provides good instructions on how to carry out a listing operation (see Sections 1.10, 1.11, and Chapter 2).

7.1 Listing form

Interviewers will have to fill a listing form for each cluster. The listing form should have four sections: (i) a header section, with general information about the cluster; (ii) a list, where each household will occupy a row; (iii) a location map of the cluster, indicating how to find the cluster; and (iv) a sketch map of the cluster, indicating the location of each dwelling, and the boundaries of the cluster.

The project team should ask to review the listing form, ideally provided in Excel format. The header section should have the following information:

- Proper identification of the cluster (that is, the cluster's unique identifier in the sampling frame of PSUs).
- Geographical information (region, province, district, locality, etc.).
- GPS information, if applicable.
- Name of interviewers.
- Date of listing operation.

The list of households should have the following columns:

- Unique serial number for the household.
- Name of the household head.
- Dwelling address or reference.
- Non-dwelling structure (non-dwelling structures, such as shops, businesses, vacant houses, etc., should be included in the list; this column is to flag these non-dwelling structures).
- Interviewer observations.

Since the list of household will span several pages, each page in the listing form should be numbered.

The location map and sketch map are simply blank pages to draw on, although it is a good idea to include a compass rose, a legend, and a space to write the cluster's unique identifier.

7.2 Listing operation calendar and staffing plan

This product consists in a calendar indicating the starting and ending dates of the listing operation fieldwork, and the number of staff required to cover all clusters. The project team should ask for the following information:

- What is the duration of the fieldwork, in weeks? Call this number D.
- Will the interviewers be organized in teams, and how? For listing operations, it is recommended to have teams of 2 interviewers, a listing expert and a mapper (the listing expert fills the list of households; the mapper draws the sketch maps).
- How many clusters is each team expected to cover per week? Call this number C. Expecting more than 3 or 4 clusters per week per team can be too ambitious.
- How many teams will there be? Call this number T.
- Is the number of teams consistent with the duration of the fieldwork? The following formula can be used to check this:

$$clusters\ in\ the\ sample = D \times C\ \times T$$

- If the result of the formula is significantly lower than the total number of clusters in the sample, then some parameter(s) must have been underestimated: the number of teams, the duration of fieldwork, or the number of clusters per week each team is expected to cover.

Also, the calendar should anticipate any forced interruptions to fieldwork, such as weather, national holidays, elections, and local festivities.

7.3 Listing operation supervision plan

The project team should review the listing supervision plan proposed by the survey firm. A good supervision plan should have the following elements:

- Revisits by a field manager to a random subsample of finished clusters, to assess the quality of the listings. A subsample of 25 percent or more of the clusters is recommended, to be selected randomly. The field manager should carry out an independent listing of 10 percent of the cluster or more.
- If problems are found in a revisit, the cluster should be relisted fully.
- If systematic problems are found in the revisits, all finished clusters should be revisited.

7.4 Listing operation manual

The project team should check that the listing operation manual contains the following:

- An introduction with the objectives of the listing operation and the structure of the manual.
- An explanation of the fieldwork calendar and staffing plan.
- Instructions pertaining confidentially and data safety.
- Instructions on how to locate clusters, and define their boundaries.
- A section explaining how to approach households, to ask their collaboration in providing basic information, such as their name and household size.
- Instructions on what the interviewer should do in front of problematic respondent reactions, such as refusal, unavailability, disinterest, distrust, etc.
- Instructions on how to fill the listing form and draw location and sketch maps.
- How to use GPS tools, if applicable.

7.5 Delivery of listing operation training

There should be 2 days of training (½ day in the office, followed by 1½ days in the field). Training in the office should include the following:

- An explanation of the training calendar and rules, including the roles and responsibilities of listing experts and mappers, the management staff, how final selection of field staff will be carried out, what candidates should do for food and transportation during training, that full attendance is required, and that tardiness is not accepted.
- A review of all the contents of the listing operation manual. All the topics listed in product 7.4 should be covered, and supported by PowerPoint presentations.
- An explanation of the field supervision procedures, the rules of fieldwork, and the consequences of not complying.
- A presentation of all the interviewer tools (listing forms, ID badges, bags, pencils, etc.)
- Exercises, where candidates must fill the listing forms and draw sketches based on predefined situations.

- Practices of GPS point captures, if applicable.

Training in the field should include the following:

- Real-life demonstrations by the instructor, of how to fill the listing form, including drawing a location map and a sketch map.
- Practices, where each interviewer is expected to list at least 50 households, and draw the corresponding sketch map. Practices of GPS point captures should also be included, if applicable.

7.6 Listing operation weekly reports and datasets

The project team should ask for weekly updates of the listing operation, which should include the list of clusters finished to date. The project team should verify that fieldwork is in schedule.

Weekly updates should also include a dataset containing all the listing forms entered onto computers to date. The dataset should replicate the structure of the listing forms, that is, there should be one row for each household, with the following variables:

- Cluster unique identifier.
- Geographical information (region, province, district, locality, etc.).
- GPS information, if applicable.
- Interviewer codes.
- Date of listing operation.
- Household unique serial number.
- Name of the household head.
- Dwelling address or reference.
- Non-dwelling structure (non-dwelling structures, such as shops, businesses, vacant houses, etc., should be included in the list; this column is to flag these non-dwelling structures).
- Interviewer observations.

The project team should review the dataset, and verify the following:

- Not more than one or two weeks have passed between the time a cluster is finished and the time it is included in the dataset.
- The number of households listed is similar to the number of households as per the latest census. If the number of households is too different, then there may be problems in the definition of boundaries, or in the quality of the fieldwork. However, the oldness and quality of the census can also explain some of these differences, and the project team should take this into account.
- Missing values are properly justified with an interviewer observation (for example, not being to obtain the household head name because of refusal).

Finally, weekly updates should include a report on the supervision activities of the survey firm, including which clusters were revisited, and the result of each revisit.

7.7 Listing operation final report and dataset

The final report and dataset should include the same elements described in the weekly reports. The final dataset of the listing operation constitutes the sampling frame of SSUs.

Additionally, the report should include the calculation of sampling weights corrected for PSU size (see Section 5.6.2).

7.8 Doing the listing operation and the survey in a single field effort

To reduce survey costs, a commonly used alternative consists in having the same team of interviewers do the listing operation and the survey in the same visit to the PSU. Although this can be a very effective cost cutting strategy, more supervision is required, as there is a perverse incentive for the teams to exclude hard-to-interview households from the listing.

8 Staffing and Training

8.1 Calendar and staffing plan

This product consists in a calendar for the survey fieldwork, basically indicating the starting and ending dates for the household interviews, and the number of staff required to cover the sample. The project team should ask the survey firm for the following information:

- What is the duration of the fieldwork, in weeks?
- Will the interviewers be organized in teams, and how? For SWIFT surveys, it is better to have the interviewers organized into teams of 2 interviewers each, plus one supervisor. Unless you need to visit rural clusters with extremely dispersed households, a third interviewer risks being idle.[6]
- How many days is each team expected to work in each cluster? For SWIFT surveys, 2 days is a reasonable time to cover all households in a cluster, regardless of the number of interviews per cluster. The number of interviews per cluster is not what drives the time estimation; increasing the opportunity of finding respondents is.
- How many teams will there be? Working with more than 5 teams (10 interviewers) should be avoided.
- Is the number of teams consistent with the duration of the fieldwork? As a rule of thumb, the following formula can be used:

$$clusters = teams \times duration\ (weeks) \times 3$$

[6] It is not recommended to use more than 2 interviewers per team, regardless of the number of households per cluster. Two interviewers should be capable of interviewing up to 20 households per cluster in 2 days (except in the case of rural clusters with extremely dispersed households). Adding more interviewers to cover a cluster in a single day is not a good option, as it decreases the opportunity of finding all respondents. In the case of rural clusters with extremely dispersed households, the simplest solution is to allow teams more than 2 days to complete these clusters. If there are many of these clusters, and time is of the essence, then adding a third interviewer to the team can allow completing these clusters in 2 days.

- If the result of the formula is significantly lower than the total number of clusters in the sample, then some parameter(s) has been underestimated: the number of teams, the duration of fieldwork, or the number of days each team is expected to work in each cluster. The formula above also assumes one day of rest per week.

8.2 Sample coverage plan

The sample coverage plan indicates which teams will visit which clusters, and in which order. It should be presented as two additional columns in the sample of selected PSUs (product 5.3). The first column should indicate which team will visit each cluster. This is normally done using a team code, so the actual names of the supervisors and interviewers are not necessary at this point.

The second column should indicate in which order each team will visit the clusters. For each team, this column should contain consecutive integer between 1 and *n*, where *n* is the total number of clusters assigned to the team.

To prevent interviewer and seasonal effects from confounding geographical comparisons, it is recommended to randomly assign teams and visit order. Randomly determining the team routes often raises concerns about travel time and cost, but there is a method, which allows randomizing (at least to certain extent) without increasing travel time and cost:

- First, estimate approximately how often teams will come back to a central location, such as the interviewer homes, the survey firm regional office, or a hotel. Also estimate the approximate number of clusters that a team will visit on each field run; we will call this number PS (package size).
- If each field run takes less than a week, then replace PS with the number of clusters that a team will visit in a week (as discussed previously, 3 clusters per week is reasonable).
- Organize the sample of selected PSUs into packages of approximately PS clusters. The clusters in each package should be chosen so as to minimize travel between them. The number of clusters in each package does not have to be exactly equal to PS; it can vary between packages, as long as the average package size is equal to PS.
- Randomly assign packages among the teams. Apply restrictions if needed. For example, if 2 teams with special language skills are required to cover some packages, then only these 2 teams should be randomly assigned among these packages. Or if 3 teams have to be permanently based in a certain region to reduce travel costs, then only these 3 teams should be randomly assigned to cover the packages in this region.
- For each team, sort the list of packages randomly, and ask teams to visit the packages in this order. Within each package, teams can visit the clusters in the order they find more convenient.

This method does not increase travel time and cost, compared to methods where clusters are not assigned randomly, but rather assigned conveniently to minimize travel. This happens because each package is designed to minimize travel, and each team only has to do

one package per field run. For each team, randomization only determines where the next run will be (a group of clusters that has been created to minimize travel), but there are no restrictions on how to behave during the run. The travel distance remains unaffected, because teams return to their starting point after each run, with or without randomization.

It is sometimes suggested to visit clusters close to headquarters first, in order to supervise teams more closely during the beginning of fieldwork. This is not acceptable, as the interviewer learning curve will be correlated with distance to headquarters. Also, it encourages poor training and unsustained supervision.

Finally, the sample coverage plan should anticipate any forced interruptions to fieldwork, such as rains and local festivities.

8.3 Supervision plan

The project team should review the supervision plan proposed by the survey firm. A good supervision plan should have the following elements:

- A description of how the survey firm will monitor the quality of the data. Monitoring should be done periodically, using programs that automatically check the consistency and exhaustiveness of the questionnaires, by field team. The way the survey management team will respond to data problems should also be specified. If data supervision is occurring remotely, much of this may depend on the connectivity that interviewers will encounter, and thus the magnitude and frequency of data upload that is possible from the field real-time.
- A description of how field supervisors will supervise the work done by interviewers. Supervisors should use three techniques, in order of importance:
    - Revisit a random subsample of households, to check the validity of some answers. It is recommended to have a subsample of 15 percent or more, to be selected randomly. The supervisor should repeat some questions, and verify that the recorded answers are true.
    - Observe some interviews directly, to evaluate the interviewer's behavior and performance. Observing about 5 percent of the interviews is recommended.
    - Reviewing that there are no inconsistencies in the CAPI questionnaire. This should be done for all interviews.

- A description of how the field manager will supervise the supervisors. This should include the following, among other things:
    - Random visits to the field, to review the work of field teams.
    - Revisits to a random subsample of households (1 percent or more is recommended) to check the validity of some answers.
    - Random telephone check-up calls, to check the validity of some answers (5 percent or more is recommended).

8.4 Interviewer and supervisor manuals

The project team should check that the interviewer manual contains the following:

- An introduction with the objectives of the survey and the structure of the manual.
- An explanation of the rules of training, interviewer selection and fieldwork.
- An explanation of the fieldwork calendar and staffing plan.
- Instructions pertaining confidentially and data safety.
- A remark about the importance of interviewing all households in the sample, and those households only.
- A remark stressing the importance of data quality, and an explanation of the supervision mechanisms.
- Instructions on the general operation of the tablet or smart phone being used.
- Instructions on the general functions of the CAPI software, such as opening, editing and backing up questionnaires, navigating through the questionnaire, recording different types of answers, etc.
- How to use GPS machines, if applicable.
- A section with good interviewing techniques, including:
    - The importance of personal appearance.
    - The importance of the introduction in obtaining the respondent's collaboration; a good manual has a short script for the interviewer introduction, between 25 and 50 words long.
    - The importance of treating people politely and respectfully at all times, and being neutral and direct at the same time.
    - The importance of reading the questions word for word.
    - How to control the interview in front of distracted respondents.
    - How to probe without changing the question meaning, and without leading the respondent. For difficult questions, specific instructions on how to probe should be provided.
    - What to do in front of respondents that don't know the response to a question, or do not wish to answer.
    - What to do in front of imprecise or complex answers. For difficult questions, specific instructions on how to deal with complex situations should be provided.
    - How to close the interview.

- Instructions on what the interviewer should do in front of problematic respondent reactions, such as refusal, unavailability, disinterest, distrust, etc.
- Instructions on which texts should be read aloud, and which should not (answer options, interviewer instructions, etc.)
- Instructions on how to detect and solve interview errors.
- Specific instructions for each question, specifically on how to deal with unclear or difficult situations.
- Rules of data management, backup and security.

Also, the project team should check there is an addendum for supervisors, containing the following:

- The rules of fieldwork, and the responsibilities of the supervisor.
- Instructions on how to locate clusters.
- Guidelines on how to organize the work of interviewers (who should do what and when).
- Instructions on how to manage the data backups.
- Instructions on how to supervise interviewers. The project team should verify that the instructions are consistent with the supervision plan.

8.5 Training materials and evaluations

The project team should review the materials for the training of interviewers and supervisors. Materials should include the following:

- A schedule of the training, consistent with the requirements set forth in product 8.7.
- A PowerPoint presentation for the introductory lecture, with the contents explained in product 8.7.
- For each section of the questionnaire, a PowerPoint presentation and a set of demonstrations and exercises, as explained in product 8.7.
- Daily evaluations, as explained in product 8.7.

Annex 4, Annex 5 and Annex 6 contain samples for the PowerPoint presentations, demonstrations and exercises, respectively.

8.6 List of interviewer and supervisor candidates

The project team should ask for a list of interviewer and supervisor candidates, containing the following information, for each interviewer:

- First and last name, age and gender
- A unique code assigned to each interviewer
- A national identification number, if available
- Education level and grade
- Current occupation status: working full time, working part time, working independently, unemployed, studying, etc.
- Previous experience working as interviewer in household surveys, measured in months in the field
- Which regions s/he is available to work in
- If any, which regions would s/he be available to work sleeping at home, without hotel per diem

The project team should review the list using the following criteria:

- Are there at least 60 percent more candidates than actually needed for fieldwork? This increases the chances of hiring better interviewers and creates competition.

- Is the gender of the applicants appropriate for the type of survey being carried out?
- Have all applicants completed secondary education?
- Are there applicants working full time, part time or independently, or are there applicants studying? Make sure that these applicants understand the conditions of fieldwork, commit to attend to all training sessions, and commit to working full time for the survey during the entire data collection period.
- Are there applicants with limited availability to travel to other regions? Prefer applicants with more availability to travel.
- Previous experience is desirable but not an essential pre-requisite for SWIFT surveys.

It is not recommended to assign interviewer and supervisor roles in advance. Ideally, all candidates should have the possibility of becoming supervisors, and the decision should be made based on the training evaluations.

If any last minute changes are made to the roster of candidates, these should be checked with the same criteria used above.

8.7 Delivery of training

It is important for the project team to participate during the training of field staff. Training is often underestimated, so the project team should oversee that it is delivered in full.

There should be at least 4 days of training, at least 2 days in the office, followed by 2 days in the field. The first day of training in the office should consist of plenary sessions, where the following topics are covered:

- An explanation of the training calendar and rules, including the roles and responsibilities of field staff, the management staff, how final selection of interviewers and supervisors will be carried out, what candidates should do for food and transportation during training, that full attendance is required, and that tardiness is not accepted.
- A review of all the contents of the interviewer manual. All the topics listed in product 8.4 should be covered, and supported by PowerPoint presentations (see sample in Annex 4).
- An explanation of the field supervision procedures, the rules of fieldwork, and the consequences of not complying.
- A presentation of all the interviewer tools (tablets, id badges, bags, pencils, etc.)
- Demonstrations in front of the class, consisting of role-playing of interviews. The instructor and an assistant follow predefined scripts, which they to convey lessons (see sample in Annex 5). The instructor can also do demonstrations by asking candidates to the front.

The second day of training in the office is for practice:

- Practice sessions of GPS point captures.
- Practice sessions, where candidates are organized in pairs, and interview each other.

- Exercises, where candidates must fill the questionnaire based on predefined situations (vignettes, see sample in Annex 6).

Training in the office should finish with a written exam. The project team should review the exam, and ensure that it meets the following criteria:

- The duration of the test is reasonable and specified in advance.
- Most of the exam consists in exercises or problems consisting in filling the questionnaire based on predefined situations.
- A few questions about general training topics are included.
- The answers to the questions are not trivial, and will actually allow discriminating which candidates are better.
- No questions have open answers. These are difficult to score without some degree of subjectivity. Except for questions to be solved on the questionnaire, all questions should be answered using true/false options, or a list of choices.
- There is a document with the exam solutions and grading criteria.

Training in the field consists in real life practices, sending interviewers to the field during 2 days. Each interviewer is expected to carry out 10 interviews. Instructors should accompany and help candidates in the field.

It can be useful to have a final meeting with selected personnel only, to answer questions from the 2 days of practice in the field.

Candidates selected as supervisors should undergo supervision training, which should include:

- The rules of fieldwork, and the responsibilities of the supervisor.
- Instructions on how to locate clusters.
- Guidelines on how to organize the work of interviewers (who should do what and when).
- Instructions on how to manage the data backups.
- Instructions on how to supervise interviewers. The project team should verify that the instructions are consistent with the supervision plan.

8.8 Selection of field staff

Final selection should be based on the exam results. The best candidates should be assigned as supervisors, and the next best as interviewers. The remaining candidates should be reserves. The project team should review the attendance lists, the exam scores, and the list of selected supervisors and interviewers.

It is difficult to grade field practices without some degree of subjectivity, but the instructors should take their observations in the field into account, and explain to the project team any suggested changes to the list of selected personnel.

9 Fieldwork

The project team should oversee fieldwork at all times, by asking the survey firm for weekly reports and datasets.

9.1 Weekly reports and datasets

Weekly reports should include the list of clusters finished to date. The project team should verify that fieldwork is being conducted in accordance with the sample coverage plan set forth in Section 8.2. The report should also include non-response rates, by cause (refusal, unavailability, etc.), and the supervision activities done by the survey firm.

Weekly datasets should include all the interviews done to date, and the project team should verify that no more than a week passes between the time a cluster is finished and the time it is included in the dataset.

Also the project team can use the weekly datasets to assess the quality of the interviews (non-response rates, inconsistencies, etc.), by field team, and ask the survey firm to take action.

The project team can also use the weekly datasets to start developing the code to generate survey indicators. This allows detecting data problems early on, and taking corrective action before the survey has ended.

9.2 Final report and dataset

Immediately following the end of fieldwork, the project team should receive the final dataset. The project team should review that the final dataset is complete (all the variables in the questionnaire are included) and properly documented (all variables and answer choices are labeled, and a description of each table is provided).

Also the project team should generate the survey indicators, and assess the quality of the data in the process. It also allows clarifying any questions on the dataset or documentation.

The project team should not rely on datasets cleaned by the survey firm. Cleaning is the process whereby errors in the dataset (such as missing and out-of-ranges values, inconsistencies, etc.) are corrected in the office, often manually. Corrections are done by imputing the values most likely to be true. When dealing with survey errors in the office, the survey firm does not have any comparative advantage over data analysts (the firm does not have additional information besides what is in the dataset already).

There are other reasons why any cleaning should be done by data analysts, and not by the survey firm. First, the ideal cleaning criteria vary depending on the type of analysis, and a survey firm would only provide one of the several possible versions of a cleaned dataset. Second, survey firms will often rely on manual cleaning, which is prone to errors and subjective imputations, and cannot be easily documented nor revised. Instead, most data

analysts rely on code to do the cleaning automatically, and are able to revise the cleaning criteria as often as needed.

The project team should also ask for a final report, which should include the following:

- A summary of the sampling design and the sampling process.
- A summary of the listing operation.
- A summary of the calendar and staffing plan.
- A summary of the sample coverage plan.
- A summary of the field staff screening, training and selection.
- A summary of the fieldwork, following the indications for the weekly reports, including the supervision done by the survey firm.
- An analysis of the fieldwork process and the quality of the dataset, which includes:
    - Non-response rates, by cause, field team, stratum, etc.
    - Average number of interviews per day per interviewer.
    - Average interview duration.
    - Percentage of missing and out-of-range values, and other errors.

10 Audit

The project team should carry out an independent audit, to verify that the answers reported by the survey are true. The audit consists in visiting and calling a subsample of households to ask a few verification questions, such as the following:

- Was your household visited by an interviewer in the last *[xx]* weeks, asking you questions about *[project]* (working for *[survey firm]*/wearing a *[blue]* T-shirt/etc.), and using a telephone to record your answers?
- Did your household respond to the interview? Who responded?
- How many people live in your household?
- Does your household own *[durable good]*?

A subsample of 100 randomly selected households is recommended. It is also recommended to distribute the subsample evenly among field teams. Finally, the audit should be carried out in parallel with the fieldwork, to detect problems in time for remedy.

Annexes

Annex 1. Sample budget

Survey parameters

Sample size	
Number of clusters	**100**
Listing operation	
Days of training	**2**
Number of teams	**5**
Average number of clusters that a team can list in 1 week	**4**
Weeks of listing	5.0
Days of listing	30
Training	
Days of training in the office	**2**
Days of training in the field	**2**
Number of candidates	14
Number of cars required	5
Fieldwork	
Number of teams	**5**
Number of interviewers per team	**2**
Total number of interviewers	10
Average number of clusters that a team can complete in 1 week	**3**
Weeks of fieldwork	6.7
Days of fieldwork	40

Survey budget template

#	Budget item	Unit price ($)	Quantity	Units	Total
1	**Fieldstaff transportation**				
1.1	Rental per vehicle	**100.00**	370	days	37,000.00
1.2	Fuel cost per vehicle	**10.00**	370	days	3,700.00
1.3	Salary and per diem for driver	**50.00**	370	days	18,500.00
	Total				59,200.00
2	**Fieldstaff salaries and per diems**				
2.1	Salary interviewer	**50.00**	776	days	38,800.00
2.2	Per diem interviewer	**50.00**	748	days	37,400.00
2.3	Salary supervisor	**100.00**	200	days	20,000.00
2.4	Per diem supervisor	**50.00**	200	days	10,000.00
	Total				106,200.00
3	**Materials**				
3.1	Hardware				0.00
3.2	Software				0.00
3.3	Office space for training				0.00
3.4	Office space for management				0.00
3.5	Training supplies				0.00
	Total				0.00
4	**Survey management**				
4.1	Survey director				0.00
4.2	Field manager				0.00
4.3	Data manager				0.00
4.4	Assistant				0.00
	Total				0.00
5	**Technical assistance**				
5.1	Study design (sampling, questionnaire, etc.)				0.00
5.2	CAPI programming				0.00
5.3	Quality control				0.00
	Total				0.00
6	**Miscellaneous**				
6.1	Printing				0.00
6.2	Communications				0.00
6.3	Insurance				0.00
	Total				0.00
Grand Total					165,400.00

Annex 2. Sample of CAPI questionnaire printout for review

Sample of CAPI questionnaire printout for review

HOUSEHOLD IDENTIFICATION

Type: Series of questions, 1 record per interview

Introduction:

ID

Label: ID Household

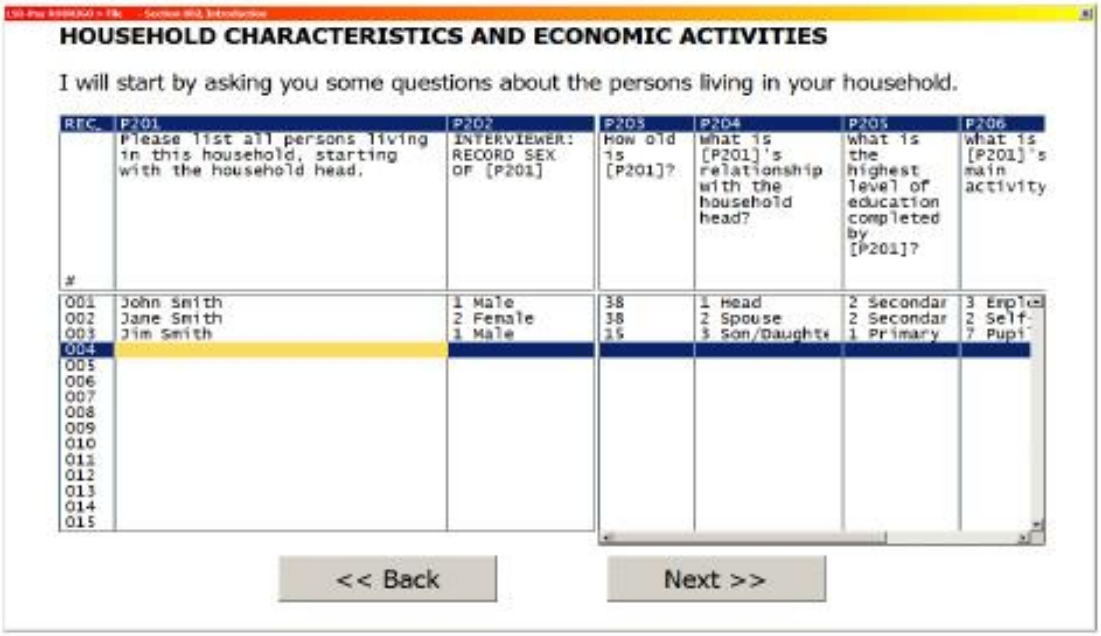
Text: INTERVIEWER: ENTER THE HOUSEHOLD IDENTIFICATION NUMBER

Answer: A number in the following range
1 - 9999

SECTION 1: HOUSEHOLD CHARACTERISTICS AND ECONOMIC ACTIVITIES

Type: Series of questions, 15 records per interview, presented in a matrix

Introduction: I will start by asking you some questions about the persons living in your household.



HOUSEHOLD CHARACTERISTICS AND ECONOMIC ACTIVITIES

I will start by asking you some questions about the persons living in your household.

REC_	P201	P202	P203	P204	P205	P206
#	Please list all persons living in this household, starting with the household head.	INTERVIEWER: RECORD SEX OF [P201]	How old is [P201]?	What is [P201]'s relationship with the household head?	What is the highest level of education completed by [P201]?	What is [P201]'s main activity
001	John Smith	1 Male	38	1 Head	2 Secondar	3 Emplo
002	Jane Smith	2 Female	38	2 Spouse	2 Secondar	2 Self-
003	Jim Smith	1 Male	15	3 Son/Daughte	1 Primary	7 Pupi
004						
005						
006						
007						
008						
009						
010						
011						
012						
013						
014						
015						

<< Back Next >>

REC_ID

Label: Record ID

Values: A number in the following range
1 - 15

P201

Label: Name
Text: Please list all persons living in this household, starting with the household head.

Answer: A text of up to 60 characters

P202

Label: Sex
Text: INTERVIEWER: RECORD SEX OF [P201]

Answer: List of options
1 Male
2 Female

P203

Label: Age
Text: How old is [P201]?

Answer: A number in the following range
0 - 99

P204

Label: Highest level of education
Text: What is the highest level of education completed by [P201]?

Answer: List of options
0 None
1 Primary
2 Secondary
3 Higher education

P205

Label: Main activity
Text: What is [P201]'s main activity?

Answer: List of options
1 Self-employed - with employees
2 Self-employed - without employees
3 Employed in the public sector
4 Employed in other sectors
5 Unpaid family worker
6 Unemployed
7 Pupil/student
8 Housewife
9 Retired
10 Person with other receipts
11 Disabled
12 Military service/prison
13 Other

P206

Label: Months worked in last 12m
Text: During the last 12 months, how many month did [P201] work, including unpaid work, vacations and sick leave?

Answer: A number in the following range
0 - 12

SECTION 2: HOUSING

Type: Series of questions, 1 record per interview

Introduction: Now I will ask you some questions about your house.

P301

Label: Sewerage
Text: Does your dwelling have sewerage?

Answer: List of options
1 Yes
2 No

P302

Label: Electricity
Text: Does your dwelling have electricity?

Answer: List of options
1 Yes
2 No

P303

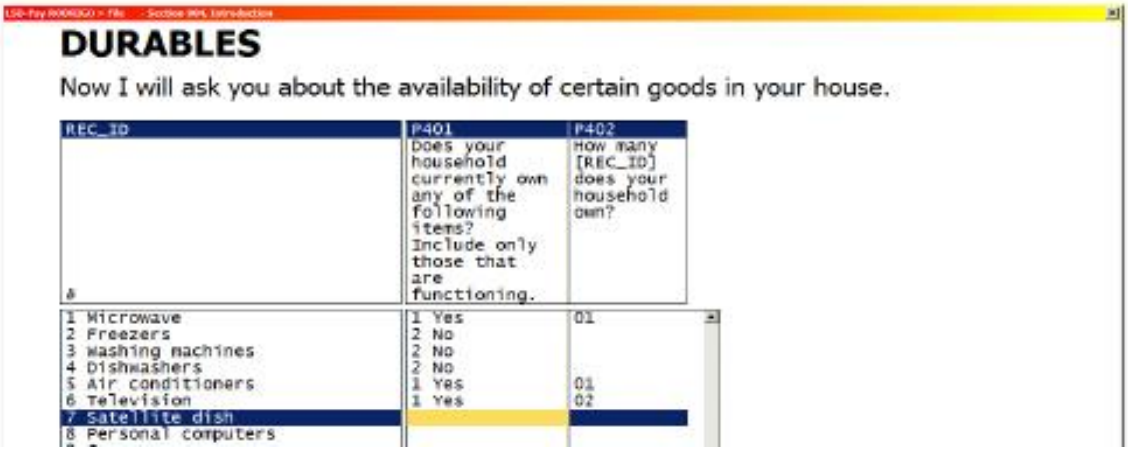
Label: Telephone
Text: Does your dwelling have landline telephone?

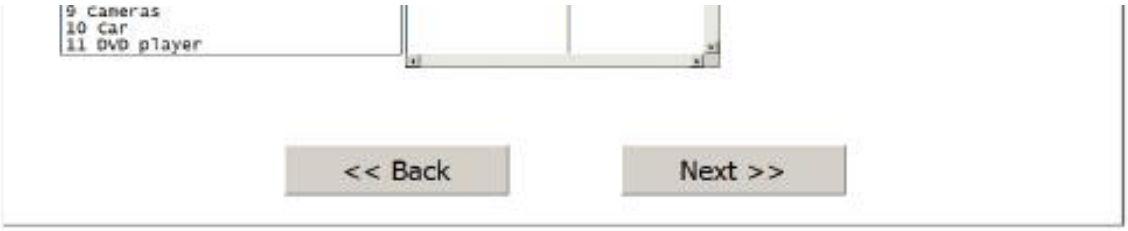
Answer: List of options
1 Yes
2 No

SECTION 3: DURABLES

Type: Series of questions, 11 records per interview, presented in a matrix

Introduction: Now I will ask you about the availability of certain goods in your house.





REC_ID

Label: Record ID

Values:

1 Microwave
2 Freezers
3 Washing machines
4 Dishwashers
5 Air conditioners
6 Television
7 Satellite dish
8 Personal computers
9 Cameras
10 Car
11 DVD player

P401

Label: Ownership

Text: Does your household currently own any of the following items? Include only those that are functioning.

Answer: List of options

1 Yes
2 No

P402

Label: Quantity

Text: How many [REC_ID] does your household own?

Answer: A number in the following range

0 - 4

Annex 3. Sample list of consistency checks

1. The first person on the list is the household head

```
If    REC_ID = 1       Then  P204 = 1
```

2. The head and spouse are of the opposite sex

```
If    P204 = 2   Then  P202 = 3 - P202(HEAD)
```

3. The household head is at least 15 years old

```
If    P204 = 1   Then  P203 ≥ 15
```

4. The age difference between the spouse and the head is 20 years or less

```
If    P204 = 2   Then  Abs(P203 - P203(HEAD)) ≤ 20
```

5. The age difference between the head and his/her children is 15 years or more

```
If    P204 = 3   Then  P203(HEAD) - P203 ≥ 15
```

6. The highest level of education completed is consistent with the age

```
If    P205 = 1   Then P203 ≥ 14
If    P205 = 2   Then P203 ≥ 17
...
```

7. If the household owns Microwaves, Freezers, Washing Machines, Dishwashers, Air Conditioners, Television, Satellite Dish or DVD Player, then the household must have electricity.

```
If    REC_ID =  1 And P401 = 1   Then  P302 = 1
If    REC_ID =  2 And P401 = 1   Then  P302 = 1
If    REC_ID =  3 And P401 = 1   Then  P302 = 1
If    REC_ID =  4 And P401 = 1   Then  P302 = 1
If    REC_ID =  5 And P401 = 1   Then  P302 = 1
If    REC_ID =  6 And P401 = 1   Then  P302 = 1
If    REC_ID =  7 And P401 = 1   Then  P302 = 1
If    REC_ID = 11 And P401 = 1   Then  P302 = 1
```

Annex 4. Sample training PPT

Introduction

Prepare slides covering at least the following topics. Refer to the manual when needed.

- About the survey
- The interviewer's job
- Rules of training, personnel selection, and fieldwork
- Training and fieldwork calendar
- Interviewing techniques
  - Establishing rapport
  - Dealing with frequent respondent questions or reactions
  - Probing
- General use of hardware and software
- ...

SWIFT
Survey of Well-being via Instant and Frequent Tracking

Day 1 - Afternoon
The questionnaire

SWIFT
Survey of Well-being via Instant and Frequent Tracking

Section 1: Household characteristics and economic activities

Insert list of topics for this section, and refer to the manual

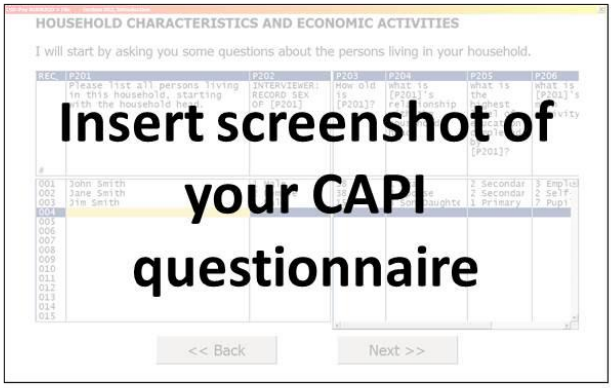
Criteria to determine who are household members and who are not.

Who is the household head.

The current educational system of grades and levels, and how to convert from old systems.

How to classify economic activities.

…



Make demonstrations of how to fill this section (see SWIFT Demonstration Sample.docx)

Section 2: Housing

Insert list of topics for this section, and refer to the manual

Waste disposal systems in the country. Which ones classify as sewerage.

What to do if electricity is not available 24 hours.

Verify by observation.

…

Make demonstrations of how to fill this section

Section 3: Durables

Insert list of topics for this section, and refer to the manual

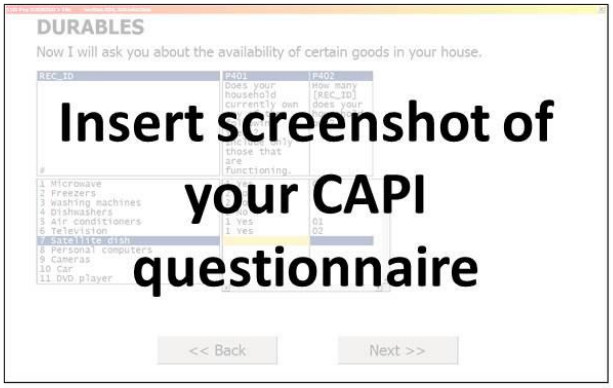
How to deal with durables used but by not owned by the household.

Include only functioning equipment.

What to do when respondents seem apprehensive about these questions.

Verify by observation if you can.

...



Make demonstrations of how to fill this section

Other sections …

Insert list of topics for other sections, and refer to the manual

...

Insert screenshots of other sections

Make demonstrations of how to fill the other sections

Annex 5. Sample training demonstration

Notes for the instructor: this demonstration requires an assistant. Your assistant plays the role of the interviewer, and reads the questions in red. The instructor plays the role of the respondent, and reads the lines in blue. As you do the demonstration, the assistant should fill in questionnaire on screen. If the instructor stops to deliver explanations, it's advisable to restart from the previous line in the script.

Materials: two copies of this guide, one for the instructor and one for the assistant. Beam the electronic questionnaire on a screen.

ASSISTANT: I will start by asking you some questions about the persons living in your household? Who is the head of this household?
INSTRUCTOR: My husband, but he's at work right now.

ASSISTANT: What is your husband's name?
INSTRUCTOR: Peter.

ASSISTANT: And his last name?
INSTRUCTOR: Dawson.

ASSISTANT: What is your full name?
INSTRUCTOR: Corinne Watson.

ASSISTANT: Now please tell me who else lives in this household.
INSTRUCTOR: Just our kids and us.

ASSISTANT: Tell me the names of all your children living in this household.
INSTRUCTOR: Katniss, Peter and Monica.

ASSISTANT: Besides your husband Peter and your children Katniss, Peter and Monica, does anybody else live in this household, even if it's not a family member?
INSTRUCTOR: No.

ASSISTANT: Do you have any babies?
INSTRUCTOR: No.

ASSISTANT: Do you have domestic workers that sleep in your house?
INSTRUCTOR: No.

ASSISTANT: How old is Peter?
INSTRUCTOR: 54.

ASSISTANT: What is the highest level of education completed by Peter?
INSTRUCTOR: Ah, he didn't study anything…dropped out of school to start working.

ASSISTANT: But he did complete primary school, didn't he?
INSTRUCTOR STOPS: This is a leading question, and is not a good way to probe. We will try again.

ASSISTANT: What is the highest level of education completed by Peter?

INSTRUCTOR: Ah, he didn't study anything…dropped out of school to start working.

ASSISTANT: In which grade did he drop out?
INSTRUCTOR: Second grade.

ASSISTANT: Of primary or secondary?
INSTRUCTOR: Secondary

ASSISTANT: What is Peter's main activity?
INSTRUCTOR: I'm sorry, what?

ASSISTANT: What is Peter's main activity?
INSTRUCTOR: Oh, we like to go hiking during the weekend.

ASSISTANT ASKS: I don't know what to do here professor.
INSTRUCTOR OPENS UP A DISCUSSION WITH THE AUDIENCE: Clearly the respondent did not understand the question. What would you do now?

INSTRUCTOR ENDS DISCUSSION: One way is to probe with simpler questions. Let's try again.

ASSISTANT: What is Peter's main activity?
INSTRUCTOR: Oh, we like to go hiking during the weekend.

ASSISTANT: OK. Does your husband work for income?
INSTRUCTOR: Yes, he's a painter.

ASSISTANT: OK. Does he work for a company?
INSTRUCTOR: No, he and his brother have their own business.

ASSISTANT: Thank you. What is your age Ms. Watson?
INSTRUCTOR: Wait, why are you asking about my husband's job?

ASSISTANT: Thank you. What is your age Ms. Watson?
INSTRUCTOR: Wait, why are you asking about my husband's job?

ASSISTANT: Employment information is very important to our study. Remember that we will use this information only for statistical purposes; so all your answers are absolutely confidential. I can answer all your questions at end of the interview. What is your age Ms. Watson?
INSTRUCTOR: I'm 48.

ASSISTANT: What is the highest level of education you completed?
INSTRUCTOR: After finishing secondary school, I went to the police academy.

ASSISTANT: Did you graduate from the police academy?
INSTRUCTOR: No, because they said I had a bad knee. It was nothing… look (SHOWS THE KNEE). Do you really think I wouldn't have made a good police officer?

ASSISTANT: Yes. The interview is going great. Let's continue. What is your main activity?
INSTRUCTOR: I take care of the kids and do everything in the house, and sell cakes.

ASSISTANT ASKS: I don't know what to do here professor.

INSTRUCTOR OPENS UP A DISCUSSION WITH THE AUDIENCE: The person has two important activities. What would you do now?

INSTRUCTOR ENDS DISCUSSION AND REFERS TO MANUAL TO EXPLAIN WHAT TO DO IN THIS CASE.

ASSISTANT: How old is Katniss?
INSTRUCTOR: 25.

ASSISTANT: What is Katniss's relationship with the household head?
INSTRUCTOR: His daughter. Katniss, Peter and Monica are our children.

ASSISTANT: What is the highest level of education completed by Katniss?
INSTRUCTOR: Secondary.

ASSISTANT: What is Katniss's main activity?
INSTRUCTOR: She studies at the university.

ASSISTANT: How old is Peter.
INSTRUCTOR: 54.

ASSISTANT: How old is your son Peter.
INSTRUCTOR: 23.

ASSISTANT: What is the highest level of education completed by your son Peter?
INSTRUCTOR: Secondary.

ASSISTANT: What is Peter's main activity?
INSTRUCTOR: He's looking for work. In the meantime he helps his father.

ASSISTANT: Does he get paid helping his father?
INSTRUCTOR: No.

ASSISTANT: How old is Monica.
INSTRUCTOR: 5.

How the screen should look at the end of the demonstration:

SECTION 1: HOUSEHOLD CHARACTERISTICS AND ECONOMIC ACTIVITIES

I will start by asking you some questions about the persons living in your household.

REC_ #	P201 Please list all persons living in this household, starting with the household head.	P202 INTERVIEWER: RECORD SEX OF [P201]	P203 How old is [P201]?	P204 What is [P201]'s relationship with the household head?	P205 What is the highest level of education completed by [P201]?	P206 What is [P201]'s main activity
001	Peter Dawson	1 Male	54	1 Head	1 Primary	2 Self-
002	Corinne Watson	2 Female	48	2 Spouse	2 Secondar	2 Self-
003	Katniss Dawson	2 Female	25	3 Son/Daughte	2 Secondar	7 Pupi
004	Peter Dawson	1 Male	23	3 Son/Daughte	2 Secondar	5 Unpa
005	Monica Dawson	2 Female	05	3 Son/Daughte		
006						
007						
008						
009						
010						
011						
012						
013						
014						
015						

<< Back Next >>

Annex 6. Sample training vignette

Materials: two copies of this guide, one for the instructor and the other for the assistant. A whiteboard marker. A tablet or phone with the SWIFT electronic questionnaire for each assistant. A copy of the statement below for each assistant.

Total duration: 30 minutes. 15 minutes to do the exercise + 15 minutes to show the solution and answer questions.

Statement: Complete section 1 based on the following situation:

In one of the addresses assigned to you, you find Marge Thompson (38 years old) with her 6-month-old baby Amelia Mendez. She says that she lives with Amelia's father, Lorenzo (40 years old), who is the head of the household. They also live with Lorenzo's father Mario, who had Lorenzo when he was 18 years old.

Both Marge and Lorenzo completed secondary school, and Mario dropped out before completing grade 2 in secondary school.

Mario is not home, visiting a friend for the weekend. He does not work and has disability pension. Lorenzo is at work, in a local restaurant. Marge stays home taking care of the baby and doing all household chores.

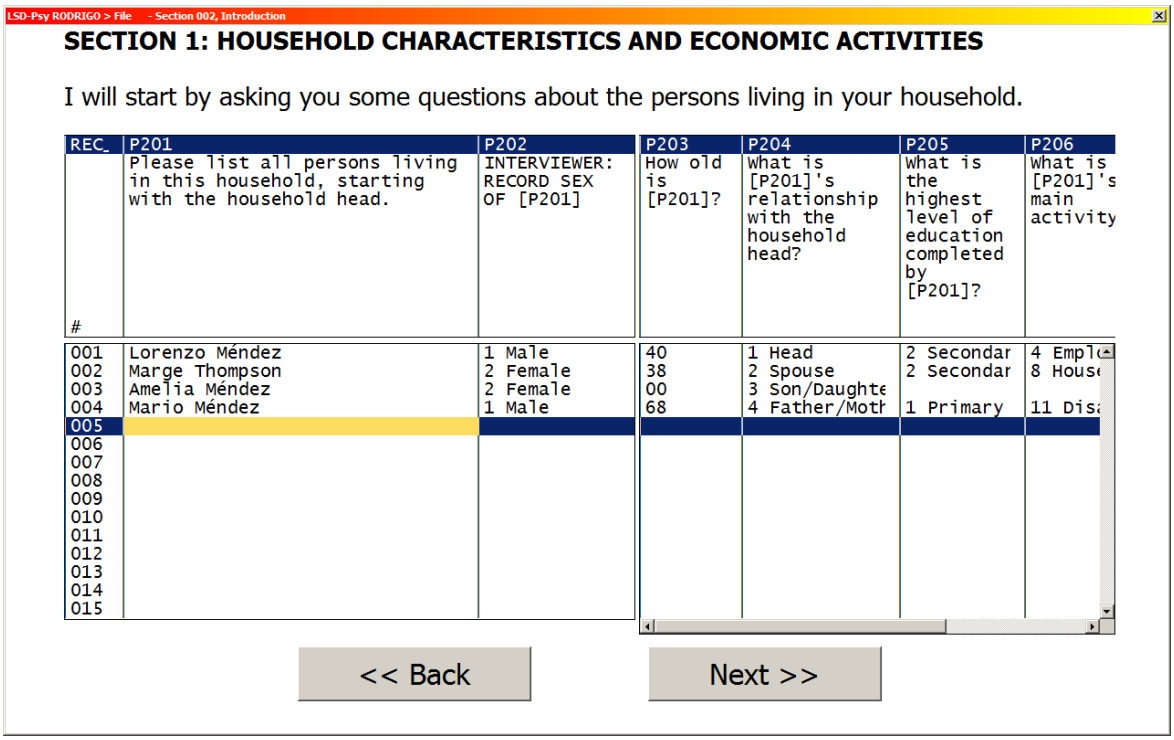
Lessons to convey:

(1) A person away from home can be a household member. Review the criteria to determine which persons qualify as household members.

(2) The highest level of education must be completed. Mario attended secondary school, but did not complete it.

(3) You must know how to classify people's occupations. Working at a restaurant is classified as "Employee in the private sector".

Solution:


LSD-Psy RODRIGO > File - Section 002, Introduction
SECTION 1: HOUSEHOLD CHARACTERISTICS AND ECONOMIC ACTIVITIES
I will start by asking you some questions about the persons living in your household.
REC_ #
P201 Please list all persons living in this household, starting with the household head.
P202 INTERVIEWER: RECORD SEX OF [P201]
P203 How old is [P201]?
P204 What is [P201]'s relationship with the household head?
P205 What is the highest level of education completed by [P201]?
P206 What is [P201]'s main activity
001 Lorenzo Méndez 1 Male 40 1 Head 2 Secondar 4 Empl
002 Marge Thompson 2 Female 38 2 Spouse 2 Secondar 8 Hous
003 Amelia Méndez 2 Female 00 3 Son/Daughte
004 Mario Méndez 1 Male 68 4 Father/Moth 1 Primary 11 Dis
005
006
007
008
009
010
011
012
013
014
015
<< Back
Next >>

Annex 7. Survey CTO vs. World Bank CAPI

A CAPI comparison: SurveyCTO versus World Bank CAPI

April 3, 2014

Alexander Skinner
Catherine Lee

SWIFT
Survey of Well-being via
Instant and Frequent Tracking

SurveyCTO

Stage 1 – Questionnaire Creation

- Questionnaire is created using excel and pre-programmed codes for variable types, name, question wording, constraints, conditionalities, etc.

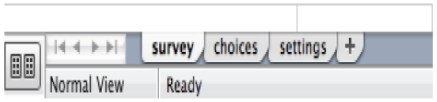
	A	B	C	D	E	F	G	H	I	J	K
1	type	name	label	hint	default	appearance	constraint	constraint message	relevance	disabled	required
2	start	starttime									
3	end	endtime									
4	deviceid	deviceid									
5	subscriberid	subscriberid									
6	simserial	simid									
7	phonenumber	devicephonenum									
8											
9	integer	numfamily	How many family members are in your household?				.>0	Please enter a number greater than zero			yes
10	note	repeatintro	Now I will ask you a few questions about each family member living with you.								
11	begin repeat	fam_group	Family Member								
12	text	name	What is your name?								yes
13	integer	months_present	Write the number of months present in household during the last 12 months				.>=0 and .<=12	Please enter a number between 0 and 12			yes
14	select_one absence	household_absence	Why was ${name} absent from the household?						${months_present}<12		
15	select_one sex	sex	What is ${name}'s sex?								yes
16	integer	age	How old is ${name}?				.<130	Please enter a valid age to continue.			yes
17	select_one highested	education	What is the highest level of education completed by ${name}?						${age}>=15		yes
18	select_one mainact	main_activity	What is ${name}'s main activity?						${age}>=15		yes
19	integer	months_worked	How many months did ${name} work? Include unpaid work, vacation and sick leave						${age}>=15		yes
20	end repeat	fam_group									
21	select_one mainincome	main_income_source	What is the main source of income for your household?								yes

survey / choices / settings

Stage 1 - 3 Sheets

- Your standard survey template will contain three sheets.
- These are:
- 1. Survey 2. Choices 3. Settings

When you open this sample workbook, you will see that it contains three worksheets: *survey*, *choices*, and *settings*:



Stage 1 - Integer

- Examples include “Age” or “Number of Family Members”
- “Age” is a simple numerical input
- “Number of Family Members” is, typically, one of the first things you seek to establish in a survey is the number of family members.
- This can be looped using a “repeat” command.

type	name	label
integer	age	What is your age?

This is the Android interface for the integer question:

Stage 1 – Repeat Command

begin repeat	fam_group	Family Member
text	name	What is your name?
integer	months_present	Write the number of months present in household during the last 12 months
select_one absence	household_absence	Why was ${name} absent from the household?
select_one sex	sex	What is ${name}'s sex?
integer	age	How old is ${name}?
select_one highested	education	What is the highest level of education completed by ${name}?
select_one mainact	main_activity	What is ${name}'s main activity?
integer	months_worked	How many months did ${name} work? Include unpaid work, vacation and sick leave
end repeat	fam_group	

Stage 1 - Constraints

- The "Number of Family Members", for example has to be at least one for the survey to make any sense.
- To prevent any accidental "0" or negative number, we can build in a constraint
- In this case, the constraint is .>0

type	name	label	hint	default	appearance	constraint	constraint message	relevance	disabled	required
start	starttime									
end	endtime									
deviceid	deviceid									
subscriberid	subscriberid									
simserial	simid									
phonenumber	devicephonenum									
Integer	numfamily	How many family members are in your household?				.>0	Please enter a number greater than zero			yes

Stage 1 – Multiple Choice

- "Gender" or consent-based questions are good examples.

The next field is a multiple choice field, which uses the select_one field type:

type	name	label
select_one yesno	consent	Would you like to continue?

The possible answers to this multiple choice question are listed on the *choices* sheet:

list_name	name	label
yesno	1	Yes
yesno	0	No

For each option in the "yesno" list, there is a *label* (which will appear to users) and a *name* (which is the associated value that will appear in the data). This is how this multiple choice question will appear to users:

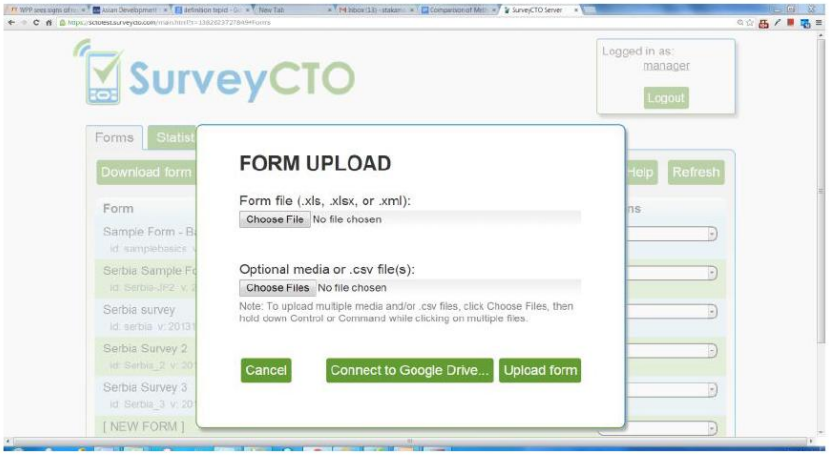
Stage 1 – Language

- Multiple language versions can be created in the excel file.

label:azerbaijani	constraint message:azerbaijani	label:russian	constraint message:russian
Nümunə şəklində xoş gəlmisiniz. Davam irəli çalmak edin.		Добро пожаловать в виде образца. Пожалуйста, проведите жду продолжения.	
Davam etmək istəyirsiniz?		Хотите продолжить?	
Sizin adınız nədir?		Как тебя зовут?	
Neçə yaşınız var?	Davam etmək üçün bir etibarlı yaş daxil edin	Сколько вам лет?	Пожалуйста, введите действительный возраст, чтобы продолжить.
Sizin adı ${name} və yaş deyil ${age}. Sağ ol.		Ваше имя ${name} и ваш возраст ${age}. Спасибо.	

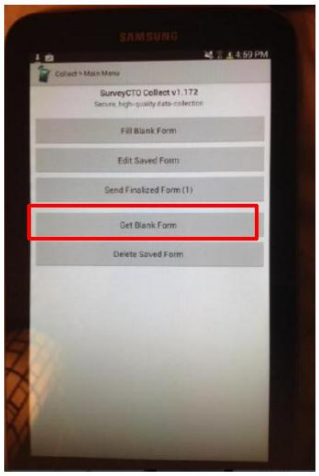
Stage 1- (cont'd)

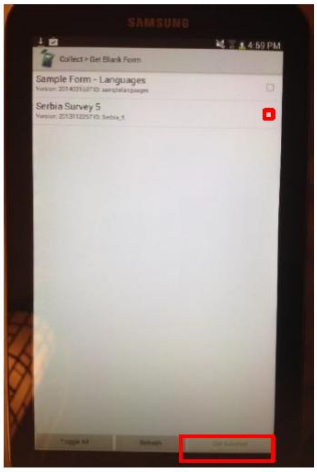
- Upload questionnaire at www.surveycto.com



Stage 2 – Data Entry

- Download your form by selecting “Get Blank Form”
- Then tick the box next to your form and click “Get Selected”





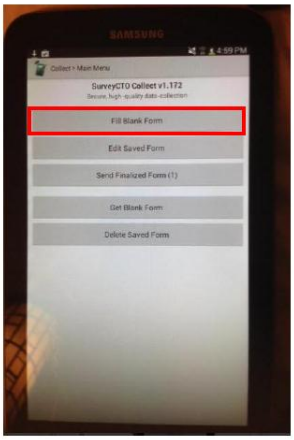
Stage 2 – (cont’d)

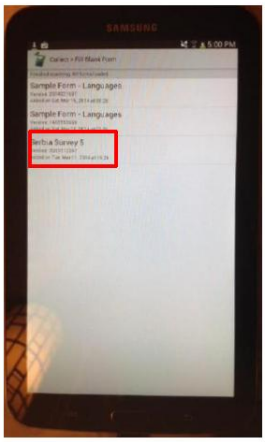
- You can download the form using an **Android-based…**

Mobile Phone OR Tablet

Stage 2 – (cont'd)

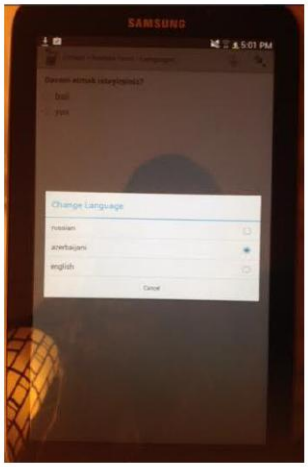
- Conduct the survey by clicking "fill blank form" and selecting your form.





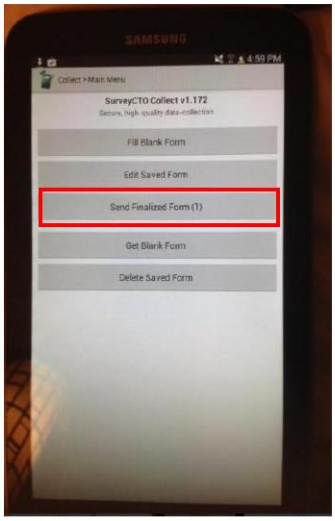
Stage 2 – (cont'd)

- The language of the questionnaire can be switched by clicking "settings" and then "language settings"



Stage 2 – (cont'd)

- Send your data to the server using "send finalised form"



Stage 3 – Download & Calculation

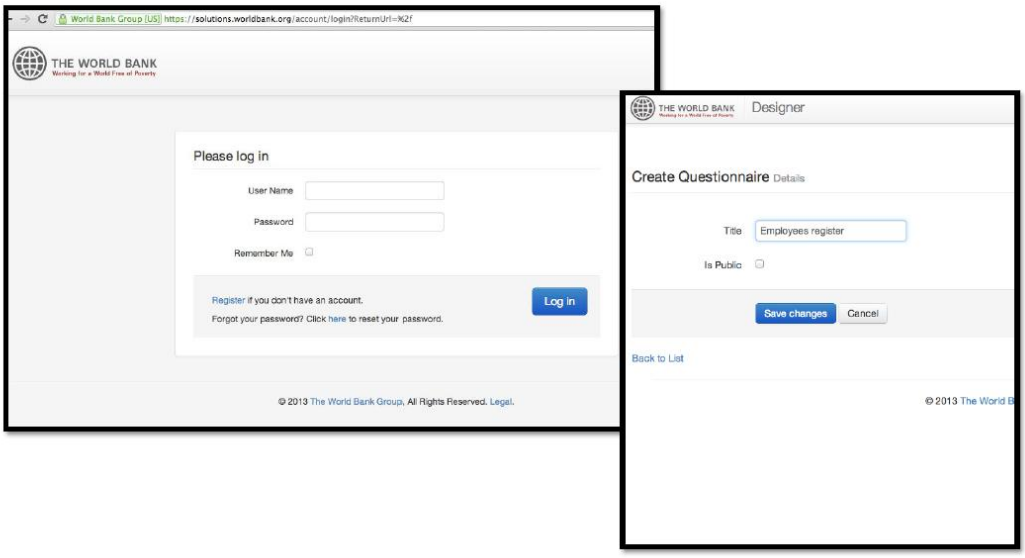
- Download the data to any computer using SCTO Client. This software can be found at www.surveycto.com

THE WORLD BANK

DEVELOPMENT RESEARCH Research at the World Bank

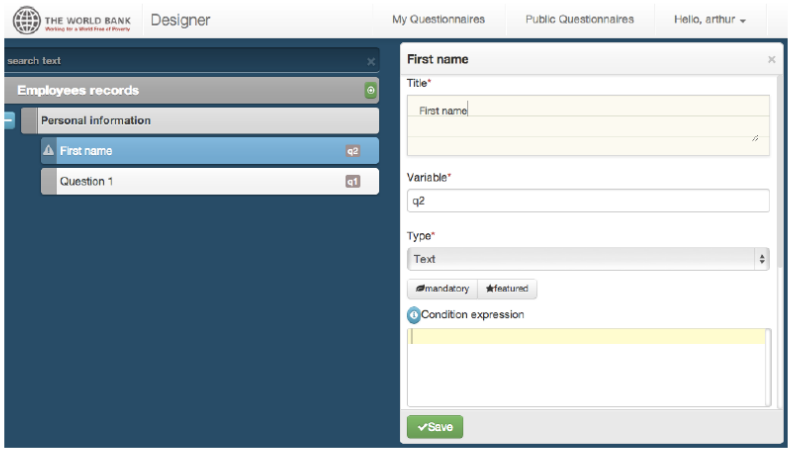
Stage 1 – Questionnaire Creation

- Create a questionnaire using Questionnaire Designer web interface at solutions.worldbank.org (i.e. no third party software involved)



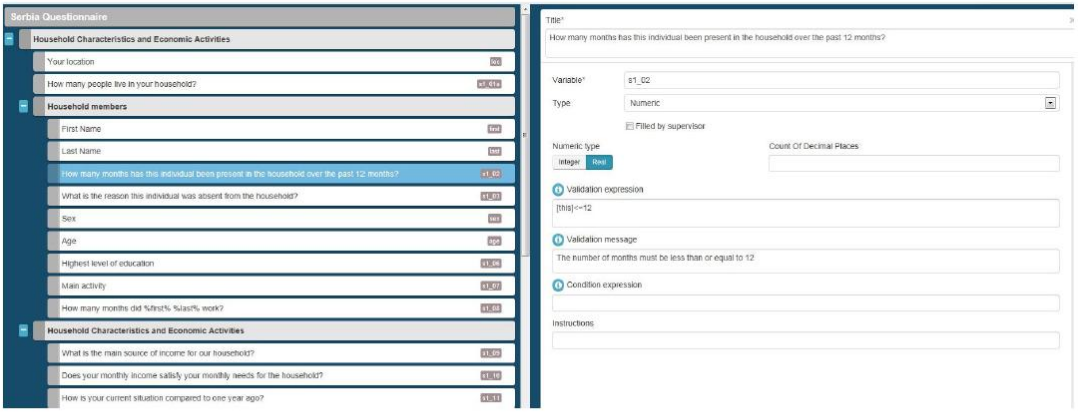
Stage 1 – (cont'd)

- Each question is created (or edited) in an interface embedded into the Designer webpage. It is visually intuitive.



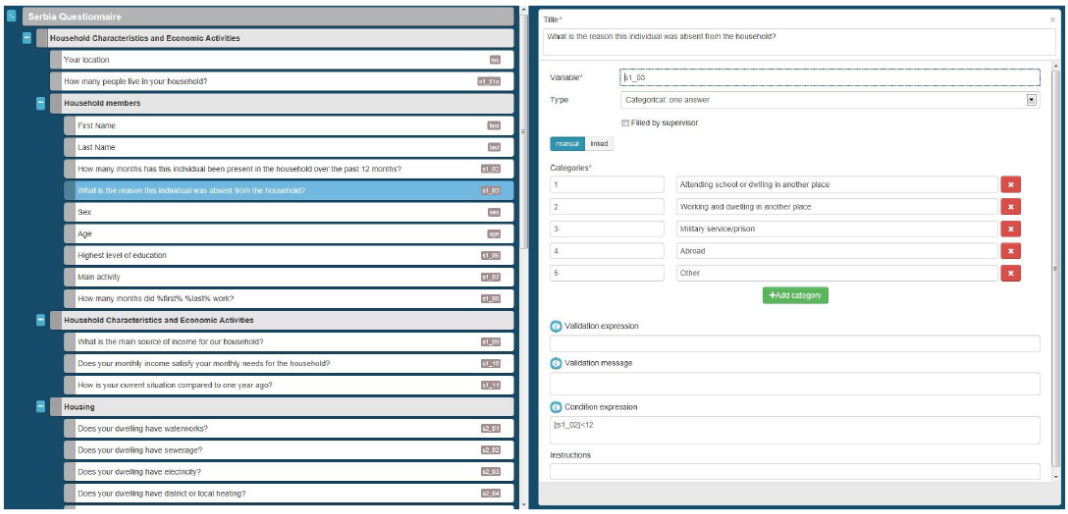
Stage 1 – (cont'd)

- Each question can take fields such as title (the text of the question), variable name, type, validation expression, validation message, conditional expression, and instructions.
- See example of numeric question with validation expression below



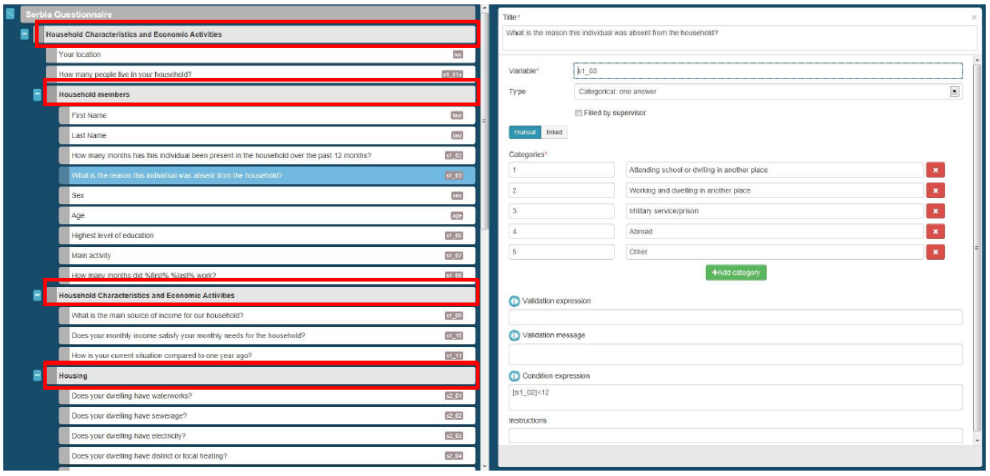
Stage 1 – (cont'd)

- See example of a multiple choice, conditional question below
- Validation or condition expressions can take operators such as +, -, *, /, ==, =!, >=, <=, and, or, contains, in



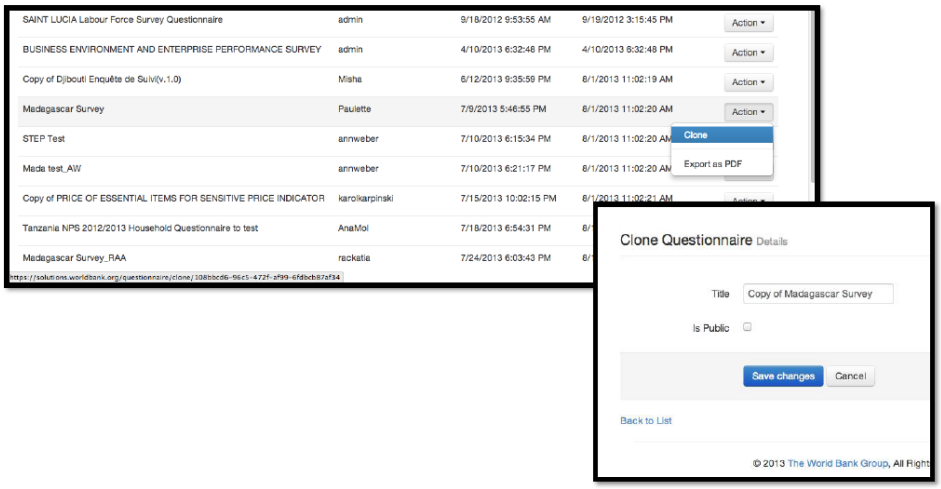
Stage 1 – (cont'd)

- Questions can be grouped under different chapters



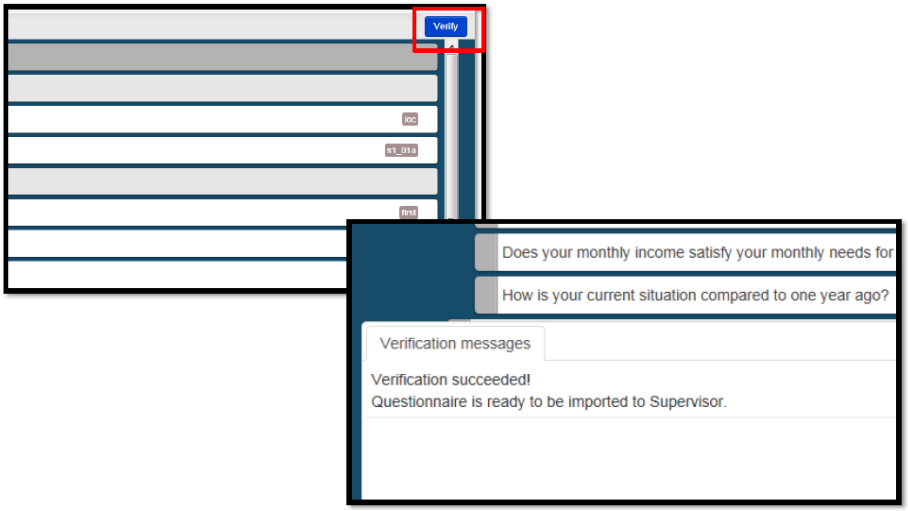
Stage 1 – (cont'd)

- To create multiple language versions of the same questionnaire, one can "Clone" an existing questionnaire (which is in, say, English) and edit the "Title" fields, i.e. the text of the questions, to a new language



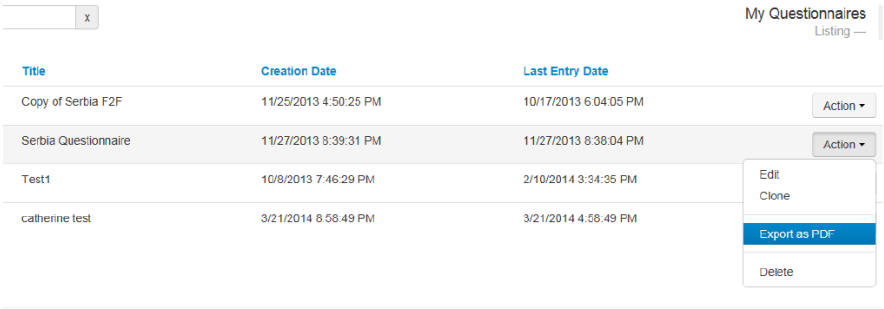
Stage 1 – (cont'd)

- At any point the designer can run the questionnaire compiler to check that the expressions entered are correct



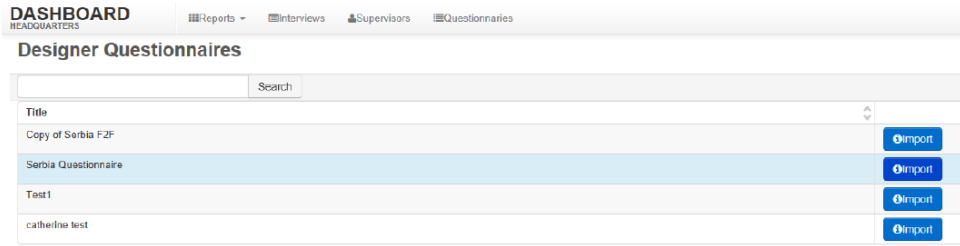
Stage 1 – (cont'd)

- The questionnaire can be exported as a pdf file, if a paper version survey is also desired.



Stage 2- Questionnaire Assignment

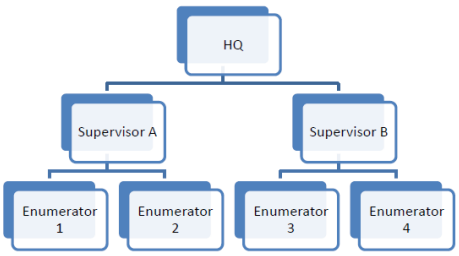
- Once the questionnaire creation is complete, we log onto swift.wbcapi.org and import the questionnaire. All survey implementation (as opposed to design) is conducted from swift.wbcapi.org (as opposed to solutions.worldbank.org)



Stage 2- (cont'd)

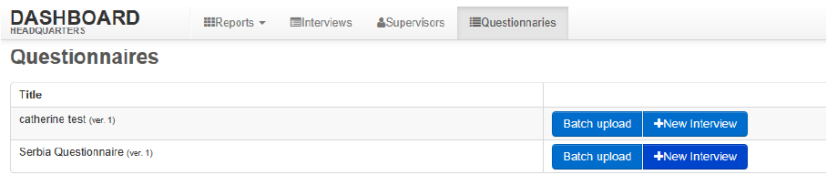
- The HQ user can then "assign" the questionnaire to a supervisor(s), who in turn assigns the questionnaire to an enumerator(s)
- An interview can be re-assigned

- *NOTE*: an 'exit-poll'-style mode, where the enumerator can start an interview on the spot without having it pre-assigned to him/her, will soon be available. However, utilizing the supervisor-enumerator oversight relationship is likely to be valuable in most of our surveys (more later)



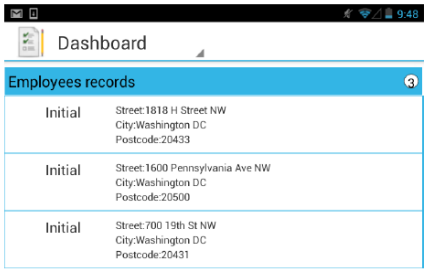
Stage 2- (cont'd)

- "Assigning" a questionnaire from HQ to a supervisor, or from a supervisor to an enumerator, is done by simply clicking in the following screen



Stage 2- (cont'd)

- Enumerators can be directed to specific household locations when they are assigned an interview. This is done via "featured" questions during the questionnaire design stage. These questions are "answered" by the individual designing the sampling frame at HQ, and can indicate, for instance, the address of the specific household(s) the enumerator is to visit. This information is displayed on the Dashboard screen of the enumerator's CAPI app on the tablet and does not appear in the actual survey.
- (To create a featured question, simply click the Featured option button while creating/editing a question.)



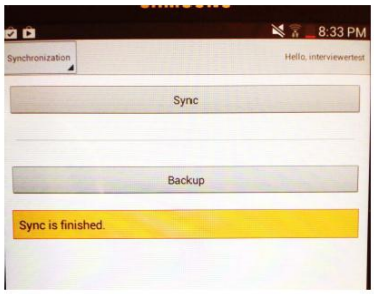
Stage 3 – Data Entry

- You can download the form using an **Android-based…**

Mobile Phone OR Tablet

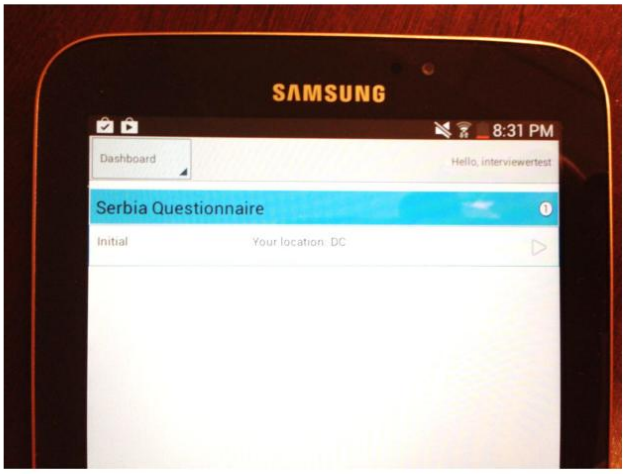
Stage 3 – (cont'd)

- Sync the CAPI app on the tablet for the latest questionnaire and assignments



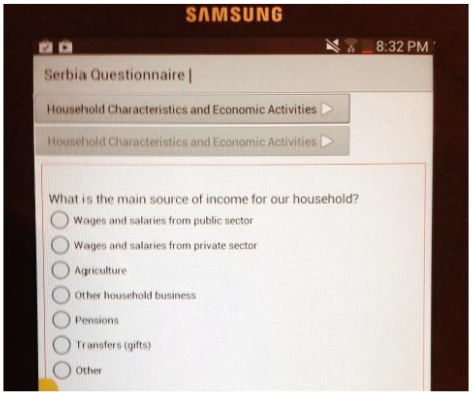
Stage 3 – (cont'd)

- Conduct the survey by clicking on a pending interview



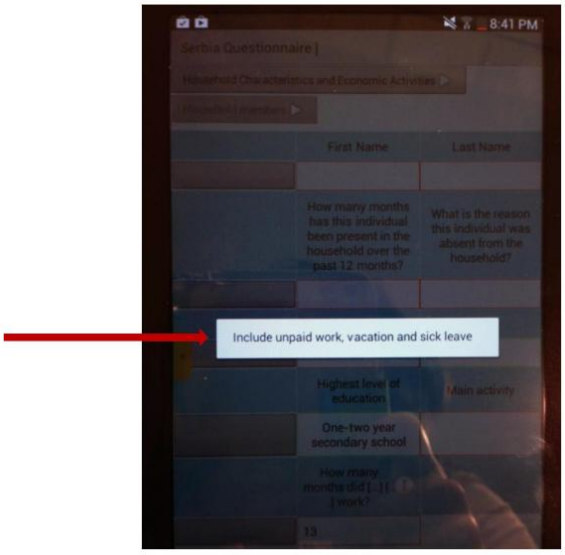
Stage 3 – (cont'd)

- Answer questions (click each chapter heading to view individual questions)



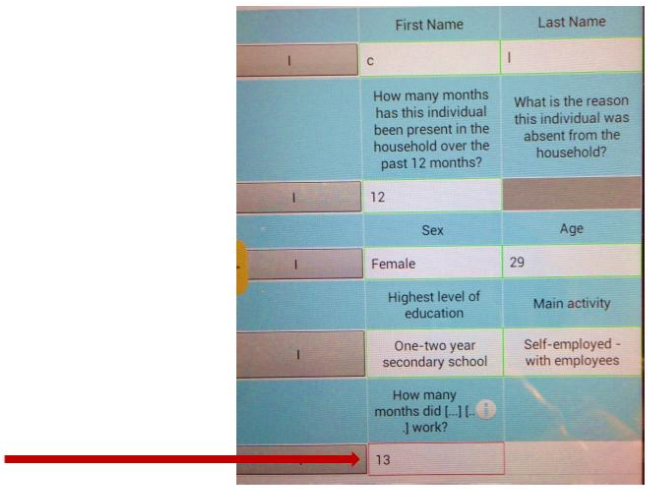
Stage 3 – (cont'd)

- If a question includes instructions, an icon with an "i" can be clicked for viewing



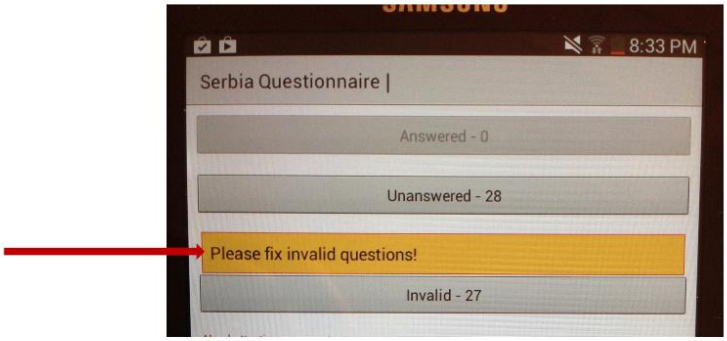
Stage 3 – (cont'd)

- If a question is answered in a way that defies the validation expression, or if a mandatory question is left unanswered, that particular response field will appear with a red border (while a valid response appears with a green border)



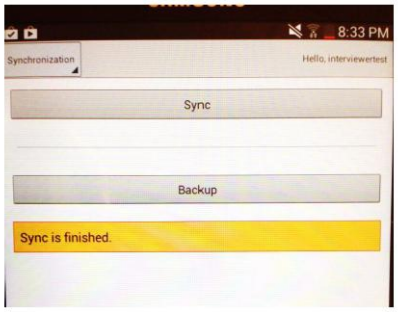
Stage 3 – (cont'd)

- If any question's response remains invalid, or empty when mandatory, at the time of an attempted survey submission by the enumerator, the app will display this screen:



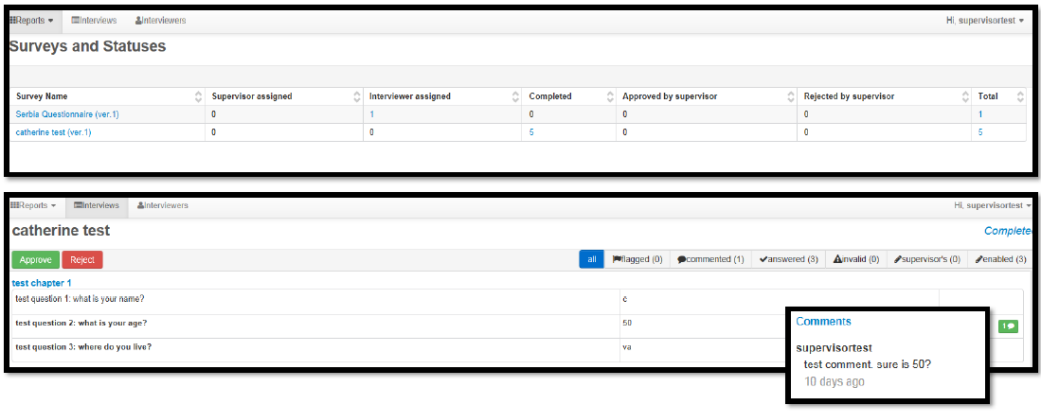
Stage 3 – (cont'd)

- Sync again to push new data (and pull any new/revised questionnaires and assignments)



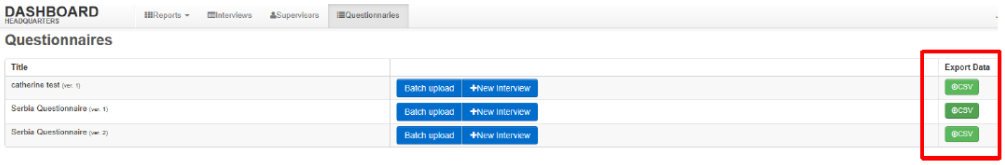
Stage 4 – Supervisor Review

- The supervisor can view the overall progress of enumerator(s) on the Dashboard, or of a single survey and can take the following actions on submitted surveys:
  - Approve, so that the data is uploaded to the HQ server
  - Comment on specific questions for the enumerator to see
  - Reject, and either make the enumerator re-conduct the survey or re-assign to a different enumerator



Stage 5 – Download & Calculation

- Download the data to any computer from swift.wbcapi.org as a .csv file, with an accompanying .do file for insheeting and labeling the raw data in Stata



Commonalities

- User-friendly & intuitive
- Responsive support teams
- GPS coordinates can be captured in a question
- Enumerator can access the necessary app in Android mobile phones and tablets

Annex 8. STATA Stepwise Regression

A procedure in the stepwise regression using STATA's command: stepwise, pr(0.051) pe(0.050): reg lnr02_pc `model'

- First, STATA runs a regression of lnr02_pc (log of per capita household expenditure at 2002 prices) on all variables included in local macro `model'
- Second, STATA removes a variable whose coefficient's significance level is lowest and lower than 5.1 percent
- Third, STATA runs a regression with the remaining variables and removes a variable whose coefficient's significance level is lowest and lower than 5.1 percent
- Fourth, STATA includes a variable in a regression model that was excluded before but whose coefficient's significance level is highest among all excluded variables and higher than 5 percent
- Repeat this procedure till neither exclusion nor inclusion can be done

References

Elbers, C., J. O. Lanjouw, and P. Lanjouw. 2002. "Micro-Level Estimation of Welfare." World Bank Policy Research Working Paper Series No. 2911, Washington, DC.

———. 2003. "Micro-Level Estimation of Poverty and Inequality." *Econometrica* 71 (1): 355–64.

James, G., D. Witten, T. Hastie, and R. Tibshirani. 2013. *An Introduction to Statistical Learning with Applications in R*. Springer Science+Business Media. New York.

Lanjouw, P., and Q. Zhao. 2014. *Using PovMap2: A User Guide*. http://go.worldbank.org/QG9L6V7P20.

Rubin, D. B. 1987. *Multiple Imputation for Nonresponse in Surveys*. Wiley. New York.

Schafer, J. 1999. "Multiple Imputation: a Primer." Statistical Methods in Medical Research 8: 3-15.

Sherpa, M., S. Radaelli, N. Yoshida. 2015. Projection of poverty into NRVA 2013. Mimeo.

STATA Corp LP. 2013. *STATA Multiple Imputation Reference Manual Release 13*. Stata Press Publication. College Station, Texas.