Weighting Procedures for Dual Frame Telephone Surveys: A Case Study in Egypt

Survey Methods: Insights from the Field, Special issue: Weighting: Practical Issues and 'How to' Approach

Mahmoud A. Elkasabi

How to cite this article : Elkasabi, M. A. (2015), Weighting Procedures for Dual Frame Telephone Surveys: A Case Study in Egypt. Survey Insights: Methods from the Field. Weighting: Practical Issues and 'How to' Approach. Retrieved from https://surveyinsights.org/?p=5291

DOI : 10.13094/SMIF-2015-00005

Abstract : Although it is a quick and non-expensive tool used to collect survey data in Egypt, the landline telephone surveys cannot reach the non-landline households, which makes up 73.4 percent of the households in Egypt according to the 2012/2013 Egypt – Household Income, Expenditure, and Consumption Survey (HIECS). Therefore, among other centers, the Public Opinion Poll Center (POPC) adopted the dual frame telephone surveys as an alternative to the landline telephone surveys, in which the landline sample is supplemented by a Pseudo-Random-Digit-Dial (Pseudo-RDD) sample of cell phones. The cell phone sample can reach the cell-only households (households with no landline but are accessible by cell phone), about 66.7 percent, based on the 2012/2013 Egypt HIECS data; this contributes to reducing the potential coverage bias due to not covering these households in the landline telephone surveys. Although both are telephone samples, different weighting procedures may apply for each sample. Moreover, the overlapping between the two sampling frame should be properly identified and adjusted in the weighting procedures. In this paper, sampling design and weighting procedures for the dual frame telephone surveys in Egypt will be discussed. Data from the Current Issues Survey, conducted by the POPC in October 2014, will be used as a case study to illustrate the weighting procedures and to support our discussions.

Introduction

The rapid development in the telecommunications resulted in a considerable accelerated change in the telephone service structure all over the world. As a result, the percentages of households with only cell phones, cell-only households, have increased in many countries; either as a result of the lifestyle in European and North American countries or the rapid development in the cell phone networks infrastructure relative to the landline phone's in some developing countries. For example, in Finland, the percentage of the cell-only households increased from about 20 percent in 1996 to 74 percent in 2009. Similarly, in Austria, 45 percent of the households in 2009 were cell-only households relative to 4.3 percent in 2000. In Portugal, the percentage of the cell-only households increased from about 7.6 percent in 2000 to 42 percent in 2009. Moreover, in 2009, more than 50 percent of the households in both Slovakia and Latvia were cell-only households, whereas in Lithuania and Czech Republic, in the same year, the percentages were 59 percent and 74 percent, respectively (Kuusela, Callegaro, & Vehovar, 2008; Mohorko, de Leeuw, & Hox, 2013). Although, we did not find more recent information about the cell-only households, the previous information indicates the constant increase in the cell-only households in these countries.

In the U.S., the percentage of the cell-only households increased from 26.6 percent in 2010 to 39.4 percent in 2014 (Blumberg & Luke, 2014; 2010). This implies a considerable coverage problem of the general population by the landline telephone surveys, which do not cover the cell-only and the non-telephone households (households without access to any telephone service, either landline or cell phones). Thus many survey research centers have been encouraged to use dual frame telephone surveys, in which two samples are used, landline sample and cell phone sample (Brick et al., 2007; Keeter, Kennedy, Clark, Tompson, & Mokrzycki, 2007; Kennedy, 2007; Link, Pattaglia, Frankel, Osborn, & Mokdad, 2007).

According to the 2005 Egypt Demographic and Health Survey (EDHS), in 2005, the landline telephone surveys used to cover about 56 percent of the residential households in Egypt, households with access to landline telephones (landline-households), as reported in (Elkasabi, 2008); The percentage has decreased to about 49.5 percent in 2008 (according to the 2008 EDHS results retrieved from the DHS-Statcompiler). According to the 2012/2013 Egypt HIECS (OAMDI, 2014), the percentage of the landline-households has dramatically decreased to 26.6 percent, whereas about 91.3 percent of the households in Egypt have access to cell phones (cell-households) in the same year. These percentages increase in the urban areas where 39.3 percent are landline-households and 94.4 percent are cell-households, and decrease in the rural areas where 16.3 percent are landline-households and 88.8 percent are cell-households. Meanwhile, about 66.7 percent of the residential households in Egypt are cell-only households, where the household does not have access to landline telephones, but at least one of the family members has access to a cell phone.

The Public Opinion Poll Center (POPC) has been conducting landline telephone surveys to measure the public's opinion toward many political, social, and economic issues in Egypt since 2003. Among these surveys is the Current Issues Survey (CIS), which is a monthly landline telephone survey. The CIS has been conducted since August 2011 to measure Egyptians sentiments toward the present economic, social and political situation. A list frame landline sampling method has been used to collect the CIS data until May 2014; due to the constant increase in the number of households without landline telephones, dual frame telephone surveys, in which the standard landline samples are supplemented by Pseudo-Random-Digit-Dial (Pseudo-RDD) cell samples, have been used to collect the CIS data, among other surveys, since June 2014. The main purpose of this change is to cover all the 93.3 percent telephone households in Egypt, who either has access to landline or cell phones. The preliminary results indicated that the dual frame design is able to reach a larger proportion of young adults (18-30 years old) and households who live in rural areas. This article describes the CIS dual frame sampling design and the weighting procedures; compares between the demographic distributions of respondents from the two samples; and presents the demographic weighted distribution of respondents from the dual frame sample and compares it with the population distribution.

Sampling Frames for Dual Frame Telephone Surveys in Egypt

The dual frame telephone surveys adopted by the POPC, including the CIS, use two frames, a list frame F_{LL}, which includes all the residential landline phone numbers in Egypt and a Pseudo-RDD cell phones frame F_C, in which a random-digit-dialing procedure is adopted to guarantee that all the cell phone numbers in Egypt has the same chance to be selected in the sample. The cell phone numbers in Egypt are composed of a three-digit provider code and an eight-digit phone number. In the adopted Pseudo-RDD procedures, all the combinations of the eight digits have the same probability to be included in the sample. Procedures for identifying and excluding the non-working prefixes, as in the standard RDD procedures, has not been set up yet, at least at this early stage (For more details about the standard RDD sampling, see Casady, & Lepkowski, 1993; Tucker, Lepkowski, & Piekarski, 2002; Waksberg, 1978).

The dual frame telephone surveys in the POPC follow a common overlapping design, in which $F_{LL} \cap F_C = F_{LC}$ as in Figure 1, where F_{LC} includes the users/households with access to both landline and cell phones, *the dual users/households*. The list frame F_{LL} covers the landline-only households F_{LO} and the dual households F_{LC}, whereas the Pseudo-RDD frame F_C covers the cell-only households F_{CO} and the dual households F_{LC} through covering the dual users. Since both frames cover the dual households F_{LC}, these households have a higher chance of being selected than others, *the dual frame multiplicity*, which should be adjusted in the weighting procedures to generate unbiased results.

The dual frame multiplicity can be avoided by screening out the duplicated units from the two frames, yielding a dual frame screening design. If the duplicated unites are screened out from the cell frame, as in Figure 2, this screening design can be considered as a stratified random sample design, where F_{LL} and F_{CO} are two separate strata (AAPOR, 2010). In the dual frame telephone surveys, since the screening process can only be done during or after the data collection, the POPC adopts the overlapping design as a cost efficient design in comparison with the screening design.

Figure 1: The Dual Frame Telephone Surveys: The Overlapping Design

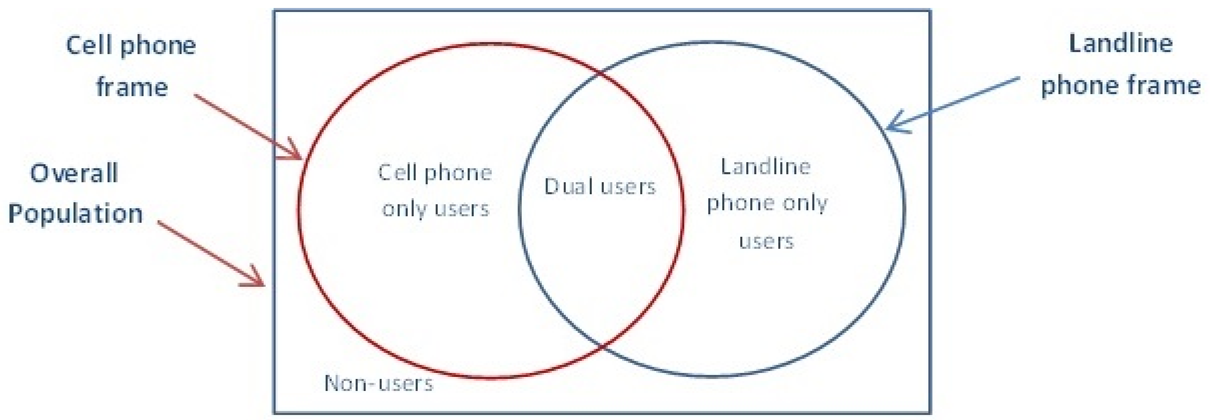
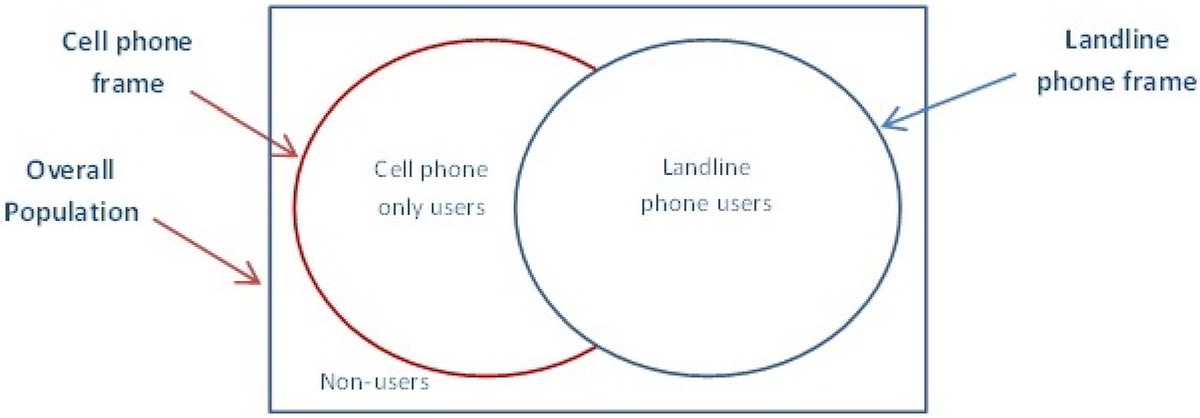


Figure 2: The Dual Frame Telephone Surveys: The Screening Design



Differences between the landline and the cell phone samples

The sample for a standard dual frame CIS is composed of two samples, a landline and a cell phone samples. The landline sample is a stratified sample by Egypt governorates selected proportionally to the number of landline phones in each governorate, selected from the list landline frame F_{LL}, which includes all the landline telephones numbers in Egypt. Whereas, the Pseudo-RDD cell phone sample is an equal-allocation stratified sample by cell phone service providers. In the CIS (October 2014), a stratified sample of about 6,965 landline numbers is selected to interview about 477 respondents, whereas a stratified sample of about 11,000 cell phone numbers is selected to interview about 1,579 respondents.

Due to the difference between the landline and the cell phones, where the former is a household owned device and the latter is a personal device, some issues should be considered in collecting the cell phone sample data. The cell phone respondent age should be at least 18 years old. This means that all respondents who are less than 18 years old are ineligible and should be screened out before the interview. For example, as indicated in Table 1, about 2.3 percent of the cell phone respondents have been screened out as ineligible respondents from the CIS (October 2014). In the landline sample, the questionnaire respondent is randomly selected from among the household adult members, whereas in the cell phone sample, the cell phone main user, who usually is the phone call receiver, is interviewed. This may lead to lower contact rate, Con R1, in the landline sample, 19.4 percent, than in the cell phone sample, 29.8 percent, as indicated in Table 1.

In contrast to other countries, in Egypt, the cell phone sample has larger response and cooperation rates and lower refusal rate than the landline sample. The percentage of the non-working, disconnected and the no-answer numbers in the landline sample, about 72.5 percent, is larger than the cell phone sample's, about 64.5 percent, as in Table 1. This indicates that the landline samples require longer time to reach the target sample size than the cell phone sample. This is because callbacks are needed to interview most of the selected respondent in the landline sample, whereas most of the cell phone respondents are interviewed in the first call.

Note that, besides the previous operational differences in the cell phone sample, more attention should be paid for the safety and the privacy of the respondent; the interviewer should explicitly ask the respondent whether he/she is in a place and a situation where they can safely answer the poll questions in a private environment. These are the relevant issues to the situation in Egypt; more details about other issues, such as legal and cost issues, in the U.S. can be found in (AAPOR, 2010; Brick et al., 2007; Link et al., 2007).

Table 1: Response Categories in the Current Issues Survey conducted by the POPC in October 2014

Response Category	The landline sample	The landline sample %	The cell phone sample	The cell phone sample %
Complete	477	6.8%	1,579	14.4%
Partial	78	1.1%	329	3.0%
Refusal and breakoff	176	2.5%	316	2.8%
Respondent never available	315	4.5%	169	1.5%
Telephone answering device	0	0.0%	208	1.9%
Always busy	239	3.4%	461	4.2%
No answer	2,477	35.5%	4,413	40.1%
Call blocking	1	0.01%	0	0.0%
Non-working number	2,519	36.2%	652	5.9%
Disconnected number	55	0.8%	2,034	18.5%
Temporarily out of service	533	7.66%	444	4.0%
Number changed	10	0.1%	117	1.1%
Non-residence	70	1.0%	24	0.2%
Regular/Vacation/Temporary residences	6	0.1%	0	0.0%
No eligible respondent	9	0.1%	254	2.3%
Total	6,965	100.0%	11,000	100.0%
Response Rate (RR1)	12.7%		21.1%	
Cooperation Rate (CoR1)	65.3%		71.0%	
Refusal Rate (RefR1)	4.7%		4.2%	
Contact Rate (ContR1)	19.4%		29.8%	

Note that RR1 = (Complete Interviews / (Complete Interviews + Partial Interviews + Refusal and break off + Non-Contacts + Other + Unknown Eligibility cases)), CoR1 = Complete Interviews / (Complete Interviews + Partial Interviews + Refusal and break off + Other) and RefR1 = Refusal and break off / (Complete Interviews + Partial Interviews + Refusal and break off + Non-Contacts + Other + Unknown Eligibility cases) and ContR1 = (Complete Interviews + Partial Interviews + Refusal and break off + Other) / (Complete Interviews + Partial Interviews + Refusal and break off + Other + Non-Contacts + Unknown Eligibility cases). These definitions are based on the American Association for Public Opinion Research (AAPOR) standard definitions (AAPOR, 2011).

The Weighting Procedures for Dual Frame Surveys in Egypt

In this section, we will illustrate the weighting procedures, which are implemented in the dual frame CIS conducted by the POPC. Generally, the weighting procedures can be classified as *main* and *supplementary* procedures. The main procedures include the design weights, in which the selection probabilities are adjusted for; it also includes the dual frame design adjustment, in which the dual frame multiplicity is adjusted. All the other weighting procedures can be classified as supplementary steps.

Let S_{LL} and S_C denote the selected samples from the landline frame, F_{LL}, and the cell phone frame, F_C, respectively, with inclusion probabilities π_i^{LL} in the landline sample and π_i^{C} in the cell phone sample. Where S_{LC} and S_{CL} denote the sample unit with access to both landline and cell phones selected in the landline sample and the cell phone sample, respectively, and S_{LO} denote the sample unit with access to landline only selected in the landline sample and S_{CO} denote the sample unit with access to cell phone only selected in the cell phone sample. Under the overlapping design, $S_{LL} = S_{LO} \bigcup S_{LC}$ and $S_C = S_{CO} \bigcup S_{LC}$. Let N_{LL} and N_C denote the frame sizes of F_{LL} and F_C, respectively, and n_{LL} and n_C denote the sample sizes of S_{LL} and S_C, respectively.

Design Weights

Base weight

This is the first step for weight computation in probability sample surveys, where every unit in frames F_{LL} and F_C, has a known non-zero probability of being randomly selected in the samples S_{LL} and S_C. In this step, each sampled element *i* (whether a respondent *r* or a non-respondent *nr*) in S_{LL} and S_C is assigned a "base weight", which is the inverse of the element selection probabilities as follows:

$$w_{1i} = \begin{cases} \frac{1}{\pi_{hi}^{LL}} & i \in S_{LL} \\ \frac{1}{\pi_{hi}^{C}} & i \in S_C \end{cases}$$

where

$$\sum_{i \in S_{LL}} w_{1i} = N_{LL}.$$

and

$$\sum_{i \in S_C} w_{1i} = N_C.$$

Regarding the landline sample, where a proportional stratified simple random sample is used, and the population is divided into H = 27 strata (governorates), where governorate h has $N_{LL,h}$ landline phones, and N_{LL} is the total number of the landline phones in the frame, which can be written as

$$N_{LL} = \sum_{h \in 1}^{H} N_{LL,h}.$$

Since a sample of size $n_{LL,h}$ landline numbers is selected in each stratum h, all numbers within each governorate have the same value of selection probability, as follows:

$$\pi_{hi}^{LL} = \frac{n_{LL,h}}{N_{LL,h}} = \frac{n_{LL}}{N_{LL}}.$$

Therefore, the base weight for the landline sample can be written as:

$$w_{1hi} = \frac{N_{LL}}{n_{LL}}. \qquad i \in S_{LL} \& h \in H.$$

$$= \frac{9,648,720}{6,965} = 1,385.3.$$

where N_{LL} = 9,648,720 is total number of landline phones in the landline list frame F_{LL}.

Regarding the cell phone sample, an equal-allocation stratified simple random sample is selected from a three-strata (service providers) population, where a sample of size $n_{C,h}$ cell phone numbers is selected from a frame of $N_{C,h}$ cell phones for provider h (all the possible combinations of the eight-digit phone number following the three-digit provider code). N_C, the total number of the cell phones in the frame, can be written as

$$N_C = \sum_{h \in 1}^{H} N_{C,h}.$$

All phone number combinations have the same selection probability, as follows

$$\pi_{hi}^{C} = \frac{n_{C,h}}{N_{C,h}} = \frac{n_C}{N_C}.$$

This because the total number of the eight-digit phone number combinations is the same for all the three providers. Therefore the base weight for the cell phone sample can be written as

$$w_{1hi} = \frac{N_C}{n_C}. \qquad i \in S_C \& h \in H.$$

$$= \frac{300,000,000}{11,000} = 27,272.7.$$

where N_c = 300,000,000 is the total number of cell phones numbers that can be generated using the Pseudo-RDD sampling procedure (cell phone numbers in F_c).

Multiplicity Adjustments

Multiplicity happens when a sampling unit has a greater probability of selection because it could have been selected through different sampling unit. Under the dual frame telephone survey design, if one household uses more than one landline telephone number for residential purposes (not solely for business, fax or computer use, etc.), the household has a greater probability of selection because it could have been selected through any of the additional telephone numbers in the household. The household weight should be adjusted to reflect the increased probability of selection. Similarly, in the cell phone sample, individuals owning or using more than one cell phone line have more chance to be selected in the cell phone sample.

To adjust for the multiplicity, more information should be collected during the interview. In the landline sample, the multiplicity information includes the number of landline telephones used to receive calls in the household, whereas in the cell phone sample, it includes the number of the cell phones owned and used to receive calls by the person. The multiplicity-adjusted weight, w_{2i}, is computed as:

$$w_{2i} = \frac{w_{1i}}{a_{1i}}.$$

where $1/a_{1i}$ is the multiplicity adjustment factor; a_{1i} is the number of landline telephones in household *i* in the landline phone sample or the number of cell phones owned or used by respondent *i* in the cell phone sample. As indicated in Table 2, most of the households own or use only one landline telephone line, 95.2 percent, whereas in the cell phone sample, 40.8 percent of the respondents own or use more than one line.

Table 2: Number of phone lines in the Current Issues Survey conducted by the POPC in October 2014

Number of phone lines	The landline sample (landline lines)	The cell phone sample (cell phone lines)
One line	95.2%	59.2%
Two lines	4.4%	33.9%
Three lines or more	0.4%	6.9%
n (respondents)	477	1579

Within Household Selection Adjustment

This step is to adjust for the "within household" selection when a household member is randomly selected from the household; this applies in case of the landline sample. The adjusted weight w_{3i} for the within household selection can be computed as the following:

$$w_{3i} = w_{2i} \times a_{2i}.$$

where a_{2i} is the number of adults (18 years or more) in household *i*; most of the households, 85.8 percent, have less than five adults, as indicated in Table 3. Regarding the cell phone sample, a_{2i} = 1 since the cell phone line is owned or used by one adult. However, if the cell phone line is owned or used by more than one adult and one is randomly selected for the survey, the number of main users can be used as the adjustment factor a_{2i} and applied to the previous equation. Note that if the survey estimates are to be reported on the household level, this step can be ignored. Also, this step does not apply, if one member is purposively selected, such as in the household head surveys, where the respondent can only be the household head or his/her partner.

Table 3: Number of adults in the landline sample in the Current Issues Survey conducted by the POPC in October 2014

Number of adults	The landline sample
One adult	16.8%
Two adults	40.7%
Three adults	14.7%
Four adults	13.4%
Five adults	7.1%
Six or more adults	7.1%
n (respondents)	477

Adjustments for non-response

In an ideal survey, all the units in the inference population are in the sample frame and all those in the sample participate in the survey. In practice, neither of these conditions occurs. Some units are not included in the frame (non-coverage) and some of the sampled units do not respond (nonresponse). The non-coverage of the cell only households is almost handled by adding the cell phone sample in a dual frame design. Regarding the non-response, it is unavoidable since we cannot contact some of the sampling units or because some sampling units refuse to cooperate in the survey. Considering that all the best practices are followed to decrease the nonresponse and to increase the response rate during the data collection, many procedures are proposed in the literature to adjust for the nonresponse.

Valliant, Dever, and Kreuter (2013) indicated that the nonresponse adjustments methods can be classified as *Weighting Class Nonresponse Adjustments* or *Propensity Score Adjustments*. The Weighting Class Nonresponse Adjustments depends on identifying weighting classes in which the response probabilities or the study variable values are homogeneous. Identifying these classes is limited with the available auxiliary variables about the sample elements, including both respondents and non-respondents. Since the values of the study variables are not available for non-respondents, a set of classes is usually identified based on response probabilities. Similarly, the Propensity Score Adjustment method depends on finding the expected response propensities based on the available auxiliary variables for both respondents and non-respondents. These response propensities can be used later to adjust the nonresponse. More details about identifying the nonresponse adjustment methods can be found in (Valliant et al., 2013).

Based on a preliminary study conducted by the POPC, the gender and the age of the sampled individuals are the two significant covariates of the response decision to the telephone surveys in Egypt. The response rate among the young age group (18 to 30 years old) tends to be lower than the response rate among the 30 + age group. Also males tend to have a lower response rate than females. Since the gender and age of the non-respondents in the landline sample are available, at least for the contacted households and using the *Weighting Class Nonresponse Adjustments* approach, gender and age are used to form four nonresponse adjustment classes, Male in 18-30 years old, Female in 18-30 years old, Male in 30+ years old, and Female in 30+ years old.

The nonresponse adjustment factor for units in class *c* (any of the four adjustments classes) can be computed using the weighted sample totals, as the following:

$$a_{3ci} = \frac{\sum_{i \in S_{c,E}} w_{3i}}{\sum_{i \in S_{c,r}} w_{3i}}.$$

where $S_{c,r}$ is the respondents set in class *c* and $S_{c,E}$ is the eligible cases set in class *c*. The nonresponse adjusted weight w_{4i} where *i* in class *c* can be written as:

$$w_{4i} = w_{3i} \times a_{3ci}.$$

The nonresponse adjustment factor for units in class *c* in the landline sample can be written as:

$$a_{3ci} = \frac{\sum_{i \in S_{LL,c,E}} w_{3i}}{\sum_{i \in S_{LL,c,r}} w_{3i}}.$$

where $S_{LL,c,r}$ and $S_{LL,c,E}$ are the respondents and the eligible sets in class *c* in the landline sample. So the nonresponse adjustment factor for the four adjustment classes in the landline sample can be calculated as in Table 4. Since there is no available information for the non-respondents in the cell sample, adjusting for the nonresponse is not possible, so $a_{3i} = 1$ would result in nonresponse unbiased estimates as long as the study variables estimates do not significantly varies between the cell phone sample respondents and non-respondents.

Table 4: Nonresponse Adjustment Factors for the landline sample in the Current Issues Survey conducted by the POPC in October 2014

Nonresponse Adjustment categories	Number of respondents (weighted numbers)	Number of non-respondents (weighted numbers)	Nonresponse Adjustment factor
Male & 18 to 30 years old	19,297	60,524	4.1
Female & 18 to 30 years old	29,823	45,612	2.5
Male & 30 + years old	148,239	207,009	2.4
Female & 30 + years old	221,043	185,080	1.8

Adjustments for Dual Frame Multiplicity

Since the landline and the cell phone frames are overlapped, this step is necessary, otherwise the survey results will be biased. Many estimators have been proposed to deal with the dual frame multiplicity, including the Fixed Weight Estimator (FWE), Hartley Estimator (HE), the Fuller-Burmeister (FB) estimator, the Pseudo-Maximum Likelihood Estimator (PML), the Pseudo-Empirical Likelihood (PEL) estimator, the Single Frame Estimator (SFE), and the Multiplicity Estimator (ME). More details about the dual frame estimators can be found in (Lohr, 2011).

Following the dual frame Fixed Weight Estimator (FWE), the dual frame adjustment factor can be written as follows:

$$a_{4i} = \begin{cases} 1 & i \in S_{LO} \\ \theta & i \in S_{LC} \\ 1-\theta & i \in S_{CL} \\ 1 & i \in S_{CO} \end{cases}.$$

where $\theta \in [0,1]$ is a composite factor used to combine the dual users from the two samples and adjust for the dual frame multiplicity. Choosing any value between 0 and 1 as the composite factor should result in unbiased survey estimators. Although the FWE estimator achieves less efficiency than other optimal dual frame estimators, it can merely be applied avoiding any complicated requirements as needed in the other optimal estimators. In the CIS, the FWE is used with 0.5 as a composite factor to combine the dual frame samples as follows

$$a_{4i} = \begin{cases} 1 & i \in S_{LO} \\ .5 & i \in S_{LC} \\ .5 & i \in S_{CL} \\ 1 & i \in S_{CO} \end{cases}.$$

The dual frame adjusted weight w_{5i} for the landline and the cell phone samples can be written as the following:

$$w_{5i} = w_{4i} \times a_{4ci}.$$

Note that the dual frame multiplicity adjustment requires information about the telephone service from the two samples; information about the cell phone usage should be collected from the landline sample respondents and information about the landline usage should be collected from the cell phone sample respondents. Consequently, besides the dual users, S_{LC} and S_{CL}, the landline only users S_{LO} and the cell phone only users S_{CO} should be identified in the landline and the cell phone sample, respectively.

Post-stratification

The Post-stratification is used to retrieve the population distribution of the demographic variables and to make the sample estimates conform to the population distribution. In the CIS, supplementing the landline sample with the cell phone sample is enough to get close to the population distribution of the age, as indicated later in Table 6, but not to the distributions of gender and place of residence. Therefore, the sample is post-stratified by region, residence area (urban/rural), and gender; the 2012/2013 Egypt HIECS is used as a source for the population distribution for these variables. Ten categories are used for post-stratification (the cross-classifications of the gender categories with three region categories and two places of residence categories). The Post-stratification Adjustment Factor for units in class *g*, in Table 5, can be computed using the weighted sample totals, as follows:

$$a_{5gi} = \frac{N_g}{\sum_{i \in r_g} w_{5i}}.$$

where r_g is the respondents in post-stratification group *g*, either from the landline sample or the cell phone sample and N_g is the population count in group *g*. The post-stratified weight w_{6i} for the landline and cell phone samples can be written as the following:

$$w_{6i} = w_{5i} \times a_{5i}.$$

Table 5: Post-stratification Adjustment Factors in the Current Issues Survey conducted by the POPC in October 2014

Gender	Region and place of residence	Sample Distribution (weighted numbers)	Population Distribution (source: the 2012-2013 HIECS)	Post-stratification Adjustment factor
Male	Urban*	14.3%	10.3%	0.72
Male	Urban Lower Egypt	12.5%	6.3%	0.50
Male	Rural Lower Egypt	17.3%	15.7%	0.90
Male	Urban Upper Egypt	10.0%	5.9%	0.60
Male	Rural Upper Egypt	11.8%	11.8%	1.00
Female	Urban*	10.9%	10.3%	0.95
Female	Urban Lower Egypt	7.3%	6.3%	0.86
Female	Rural Lower Egypt	6.6%	15.7%	2.39
Female	Urban Upper Egypt	5.2%	5.9%	1.14
Female	Rural Upper Egypt	4.1%	11.8%	2.84

*include the frontier governorates

Note that the post-stratification needs information about the population distribution across all the cross-classifications groups. In our design, the population distribution across the post-stratification variables is available. However, the more post-stratification variables used, the more information needed for the post-stratification. At the same time, having many post-stratification groups may result in large post-stratification factors for some groups, due to their small sample size. Another method used to retrieve the population distribution is the *raking*. In this method, only the marginal distributions of the raking variables are needed. More details about post-stratification and raking can be found in Valliant, et al. (2013).

Differences between the landline and the cell phone samples estimates

In this section, we explore the differences between the landline and the cell phone samples' estimates; both the weighted and the un-weighted estimates are considered. Moreover, in addition to the FWE dual frame estimates, the post-stratified FWE dual frame estimates will be added to the comparison. The deviations between the samples' estimates and the population parameters, based on the 2012/2013 Egypt HIECS, are used to support our discussion and evaluate the dual frame design performance.

Table 6 presents the un-weighted CIS survey estimates for the landline sample and the cell phone sample, in the second and the third columns, respectively. The un-weighted landline sample estimates deviate from the HIECS's population distributions. Respondents in the landline sample are more likely to be female (60 percent versus 50.9 percent in the HIECS), old (88.3 percent are 30 years old or older versus 65.6 percent in the HIECS), highly educated (39.2 percent with more than high school versus 16.6 percent in the HIECS) and to live in urban governorates (34.4 percent versus 20.7 percent in the HIECS) and in urban areas (74.8 percent versus 44.6 percent in the HIECS). On the other hand, respondents in the cell phone sample are more likely to be male (67.8 percent), younger (33.3 percent), less educated (30.8 percent with more than high school), and live in lower and upper Egypt governorates (42.6 and 30.8 percent, respectively) and in rural areas (37.8 percent) than respondents in the landline sample. Although the cell phone sample distributions of the age and the residence region are close to the population distribution, the cell phone respondents still tend to be highly educated and to live in urban areas in comparison with the population distributions.

After adjusting for the non-response, the w4-weighted estimates of the landline sample, in the fourth column, when compared with the population distribution in the first column, have somewhat improved. The weighted distribution of the gender is almost the same as the population distribution, whereas the differences between the weighted distribution of the age and the population distribution has been attenuated; by applying the weights, the landline sample represents more young respondents (21.5 percent) than the un-weighted sample. Regarding the w4-weighted estimates of the cell phone sample, in the fifth column, there are no noticeable differences between the weighted and the un-weighted distribution. The FWE dual frame w5-weighted distribution of the age, in the sixth column, matches the population distribution. Other differences in gender and place of residence have been attenuated through the post-stratification as indicated in the post-stratified w6-weighted distributions of the gender and the place of residence in the last column.

Table 6: Differences between the landline and cell phone samples estimates and the 2012/2013 Egypt HIECS Demographics (Sample estimates are based on the Current Issues Survey conducted by the POPC in October 2014)

Demographics	Population (The 2012/2013 HIECS)	Landline sample Un-weighted %	Cell phone sample Un-weighted %	Landline sample w4-weighted* %	Cell phone sample w4-weighted* %	Dual frame sample w5-weighted** %	Final dual frame sample w6-weighted*** %
Gender							
Male	49.1%	40.0%	67.8%	49.7%	66.9%	66%	50.4%
Female	50.9%	60.0%	32.2%	50.3%	33.1%	34%	49.6%
Age							
18 to 30	34.4%	11.7%	33.3%	21.5%	33.1%	33%	34.3%
30 +	65.6%	88.3%	66.7%	78.5%	66.9%	67%	65.7%

Education							
Less than high school	52.4%	23.8%	27.7%	25.9%	30.0%	32.2%	34.3%
High school	31%	37.0%	41.5%	36.1%	40.5%	41.3%	40.4%
More than high school	16.6%	39.2%	30.8%	38.0%	29.5%	26.5%	25.3%
Region							
Urban Governorates	20.7%	34.4%	26.5%	31.2%	26.4%	25%	20.6%
Lower Egypt Governorates	43.8%	40.0%	42.6%	38.8%	43.4%	43.6%	43.7%
Upper Egypt Governorates	35.5%	25.6%	30.8%	30.0%	30.2%	31.3%	35.7%
Area							
Urban	44.6%	74.8%	62.2%	67.3%	61.9%	60.1%	45.1%
Rural	55.4%	25.2%	37.8%	32.7%	38.1%	39.9%	54.9%
Telephone Service							
Landline Only users	2%	18%	NA	15.7%	NA	2.1%	2.5%
Landline and Cell phone users	24.6%	82%	31.8%	84.3%	29.7%	21.7%	20 %
Cell phone only users	66.7%	NA	68.2%	NA	70.3%	76.3%	77.5%
No phone	6.7%	NA	NA	NA	NA	NA	NA

[*] Weights accommodate sampling design, multiplicity, within household selection and non-response adjustment. The latter and the within household selection adjustments do not apply in case of the cell phone sample.

[**] FWE dual frame estimates.

[***] Post-stratified FWE dual frame estimates.

Discussion

Although the same weighting procedures apply for both the landline sample and the cell phone sample, many differences between the two samples should be considered. These differences emerged due to the difference between the landline phone as a device owned and used by the whole household members and the cell phone as a personal device. These differences can affect some of the weighting steps such as the adjustments for the multiplicity and for the within household selection. At the same time, since it requires information about the non-respondents, adjusting for the non-response may – or may not – apply in the two samples. It's worth noting that adjusting for the dual frame multiplicity is commonly overlooked in practice, which leads to biased survey estimates.

In this article, we highlighted different aspects of the POPC practice with the dual frame telephone surveys, especially in the CIS (October 2014). We focused on the weighting procedures used to calculate the sampling weights for the dual frame telephone surveys. This included the weighting procedures for the landline telephone sample and the weighting procedures for the cell phone sample. The weighting procedures included calculating the design weights, adjusting for multiplicity, adjusting for within-household selection and adjusting for the nonresponse. These procedures are performed separately for each sample. The adjustment for the dual frame multiplicity step combines the two samples' estimates to produce unbiased dual frame estimators. Finally, the post-stratification step comes to retrieve the population distribution for some demographic variables based on the combined dual frame sample.

The differences between the landline and the cell phone sample estimates support the POPC decision regarding the transition to the dual frame telephone surveys. The cell phone sample reaches more male respondents and young age respondents than reached by the landline sample. At the same time, the cell phone sample reaches more respondents who live in Lower and Upper Egypt governorates and who live in rural areas than reached by the landline sample. This means that the cell phone sample somehow reaches to the under-covered groups by the landline sample. Therefore, adding the two samples in a dual frame design should reduce the non-coverage problem in the landline sample, at least for persons who have access to phone services, specifically landline or cell phone. More research is needed to explore the operational properties of the cell phone samples in Egypt. Also, the "crude" Pseudo-RDD procedures used in Egypt so far need to be developed; a modified technique should be developed following the standard RDD sampling.

References

1. Blumberg S.J., & Luke J.V. (2014). Wireless substitution: Early release of estimates from the National Health Interview Survey, January–June 2014. *National Center for Health Statistics.* Available from: http://www.cdc.gov/nchs/nhis.htm.
2. Blumberg S.J., & Luke J.V. (2010). Wireless substitution: Early release of estimates from the National Health Interview Survey, January–June 2010. *National Center for Health Statistics*. Available from: http://www.cdc.gov/nchs/nhis.htm.
3. Brick, J. M., Brick P.D., Dipko, S., Presser, S., Tucker, C., & Yuan, Y. (2007). Cell Phone Survey Feasibility in the U.S.: Sampling and Calling Cell Numbers versus Landline Numbers. *Public Opinion Quarterly*, 71:23–39.
4. Casady, R.J. & Lepkowski, J.M. (1993). Stratified telephone survey designs. *Survey Methodology*, 19, 103-113.
5. Demographic and Health Surveys – DHS, STATcompiler. (n.d.). Retrieved December 11, 2014, from http://www.statcompiler.com/
6. Elkasabi, M. (2008). *Handling Non-coverage Bias in Telephone Surveys in Egypt* (Master's thesis). Faculty of Economics and Political Scieneces, Cairo University.
7. Keeter, S., Kennedy, C., Clark, A., Tompson, T., & Mokrzycki, M. (2007). What's Missing From National Landline RDD Surveys?: The Impact of The Growing Cell-Only Population. *Public Opinion Quarterly,* 71:772–792.
8. Kennedy, C. (2007). Evaluating the Effects of Screening for Telephone Service in Dual-frame RDD Surveys. *Public Opinion Quarterly,* 70:750–771.
9. Kuusela, V., Callegaro, M., & Vehovar, V. (2008). The Influence of Mobile Telephones on Telephone Surveys, Chapter 4 in Lepkowski, J.M., Tucker, C., Brick, J.M., de Leeuw, E.D., Japec, L., Lavrakas, P. J., Link, M. W., Sangster, R. L. (Eds.) Advances in Telephone Survey Methodology, (pp. 87-112). Wiley, NewYork.
10. Link, M.W., Battaglia, M.P., Frankel, M.R., Osborn, L., & Mokdad, A.H. (2007). Reaching The U.S Cell Phone Generation: Comparison of Cell Phone Survey Results With an Ongoing Landline Telephone Survey. *Public Opinion Quarterly,* 71:814–839.
11. Lohr, S. (2011). Alternative Survey Sample Designs: Sampling with Multiple Overlapping Frames. *Survey Methodology*, 37, 197-213.

12. Mohorko, A., de Leeuw, E., & Hox, J. (2013). Coverage Bias in European Telephone Surveys: Developments of Landline and Mobile Phone Coverage across Countries and over Time. *Survey Methods: Insights from the Field*. Retrieved from https://surveyinsights.org/?p=828
13. OAMDI, (2014). Harmonized Household Income and Expenditure Surveys (HHIES),http://www.erf.org.eg/cms.php?id=erfdataportal. Version 2.0 of Licensed Data Files; HIECS 2012/2013 – Central Agency for Public Mobilization and Statistics (CAPMAS). Egypt: Economic Research Forum (ERF).
14. The American Association for Public Opinion Research. (2010). *Cell Phone Task Force Report*. AAPOR.
15. The American Association for Public Opinion Research. (2011). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 7th edition*. AAPOR.
16. Tucker, C., Lepkowski, J.M., & Piekarski, L. (2002). The current efficiency of list-assisted telephone sampling designs. *Public Opinion Quarterly*, 66: 321-338.
17. Valliant, R., Dever, J.A., & Kreuter, F. (2013). *Practical Tools for Designing and Weighting Survey Samples*. Springer Science & Business Media.
18. Waksberg, J. (1978). Sampling methods for random digit dialing. *Journal of the American Statistical Association*, 73:40-46.