EXTENDED ABSTRACT

1. INTRODUCTION

Frame is crucial for sampling design in household surveys. The sampling frame has significant implications on the cost and the quality of any survey. It is so hard to come across perfect frame, because there may some unwanted problems in sampling frame. With the system called Address Based Population Registry System, all citizens' address information are known literally and updated continuously. In Turkey, the Address Based Population Registry System is used as sampling frame for the TDHS-2008 and 2013. In this thesis, ABPRS will be assessed and compared to the housing information which is obtained from the listing operation of TDHS-2013 and ABPRS 2012 obtained from TURKSTAT. What I want to investigate is the Address Based Registration System as sampling frame, especially in rural areas.

Hence, in the first part of the thesis, the reliability of Address Based Population Registry System(2012) will be assessed especially based on TDHS-2013 listing data. In the second part of the thesis, the usability of Address Based Registration System as sampling frame will be discussed with comparing number of occupied housing numbers; assessing the numbering of villages; evaluating the numbering of settlements which seems like province, but have village characteristics and last but not least, how many times the TURKSTAT lists was used by listing staff throughout the listing operation.

1.1 SAMPLING FRAME

A short definition of sampling frame is a source material or device from which a sample is drawn. It is a list of all those within a population that can be sampled, and may include individuals, households or institutions. Sometimes more than one set of material is needed for household surveys, due to their multi-stage nature. The first stage of the selection may be performed from area frames. For instance, households may selected either from a list frame or an area in the last stage.

Deciding on a proper frame is an important consideration based on the relationship between survey target population and unit of selection. The unit of selection determines the frame. It should be emphasized that it is also the thing that determines the probability of selection at the last stage.

As discussed above, the sampling frame should capture, in a statistical sense, the target population. Beyond that, a perfect sample frame is one that is complete, accurate and up-to-date. These are ideal properties that are unattainable in household surveys. Nevertheless it is essential to strive for them either in constructing a frame from scratch or using one that already exists. The quality of a frame may be assessed in terms of how well its idealized properties relate to the target population (Turner, 2003).

1.2 SAMPLING FRAME IN TURKEY

In Turkey, different criteria have been used to describe urban and rural settlements. In the demographic surveys of the 1970s, a population size of 2,000 was used to differentiate between urban and rural settlements. In the 1980s, the size was introduced as 10,000 and in some surveys in the 1990s as 20,000. A number of surveys used information on the administrative status of settlements in combination with population size for the purpose of differentiation.

Turkey Demographic and Health Surveys (TDHSs) can be considered as an example for the sampling frame in Turkey. For instance for the TDHS-2013, the urban frame included a list of provincial centers, district centers and other settlements with populations larger than 10,000. The rural frame contained all district centers, sub-district and villages excluded from the urban frame. Primary information on all settlements in Turkey was gained from the 2012 Address Based population Registry System (ABPRS).

In Turkey, there are no enumeration areas into which settlements are divided. Enumeration areas are small area units with well-defined boundaries. EAs are useful while conducting household surveys because they ensure the geographical proximity of selected households and reduce listing errors provided the boundaries are well defined.

In TDHS-2013, the sampling frame is obtained from the 2012 ABPRS. TURKSTAT provided settlement lists, which included the populations of provinces, districts, sub-districts and villages. In settlements whose populations were larger than 10.000, settlements divided into blocks that each include approximately 100 household. The same procedure is applied for big villages (larger than 140 household); TURKSTAT creates the blocks of around 100. For villages in which the number of households is more than 36 and less than 140, no blocks were created, the village was defined as a block itself. The smallest villages (less than 36 household), were combined with other small villages that were is in same province/district/sub-district.

2. DRAWBACKS OF ADRESS BASED POPULATION REGISTIRATION SYSTEM OF TURKSTAT AS SAMPLING FRAME

Even though no frames were provided for villages in TDHS-2008, the frames are available for TDHS 2013. TURKSTAT was not able to provide household list from the National Address Data Base for settlements without municipalities from which 132 clusters were to be drawn for the TDHS-2008. For these settlements the list of households had to be prepared in the field. In the case of small settlements(less than 250 households), the entire settlements was listed. In the case of the small number of settlements in which there were more than 250 households, an estimate of the total number of households in the settlements was obtained through a quick count, and 250 households were listed(HUIPS, 2009).

There are some drawbacks for ABPRS and the deficiencies of the ABPRS of TURKSTAT can be summarized as follows:

- Although TURKSTAT provides frames for villages, it is not satisfactory. In some cases, there is no observable house numbering. For some of the villages with observable house numbering, the numbers are either scattered or missing. However, there are two disadvantages that were observed. First one is operational difficulty. The field staff could not find addresses properly. Secondly, due to the difficulties in identifying the borders of the blocks chosen by TURKSTAT, listing either failed to cover the whole block, or covered more than the block defined in the first place. This leads to errors in selection probabilities: The numerator ends up with a different area unit than intended.

- There is no proper corresponding map for block lists provided by TURKSTAT. The logic of selecting streets for blocks is based on geographical proximity. TURKSTAT has a code in order to perform this selection. Although, streets are close to each other generally, in some cases there are big gaps between two relevant streets. This is related to lack of EAs in Turkey. İf there were EA, the selected area would become clear.

- Finally, lots of changes occur at any moment in life, like the construction/demolition of a new building or a new street. The list provided by TURKSTAT presents cross-sectional information on addresses that may not be up-to-date.

3. THE BENEFITS OF ABPRS

With this system, gender, age and education information of even the people living in the smallest residence will be known at any moment and accordingly updated information on the planning of health, education, dwelling and social services will be possessed.

Moreover, production of statistics about population size in residential base, age and sex, birth, death, marriage, divorce, migration and education will be possible in shorter periods.

4. LIMITATIONS OF ABPRS

i) In this system, some newborn babies do not get registered on time. For instance, in rural areas, some children are not registered until primary education. The gap between the birth dates to registration time turn into deficiency for ABPRS.

ii) Some people have no Turkish Republic Identification numbers. These people are not included in the system as well.

iii) In some cases, once a person dead, his/her registration information can't be erased from the system immediately. His/hers identification information may stay in system for a while.

iv) Another drawback of the system is moving. People can move one place to another, but their identification information remains in initial place because they fail to report this event.

5. DATA AND METHODS

In this study, TDHS 2013 listing data and ABPRS 2012 data will be used. According to these data; followings will be assessed;

- Numbers of occupied dwellings will be compared (as pointed out by listing and ABPRS data)
- Village numbering will be assessed.
- Some blocks defined as urban areas show village characteristics. One of the aim of this study to assess their numbering.
- Another goal of the thesis is to reveal how many of TURKSTAT lists was used or not throughout listing operation, as stated by listing staff on listing forms.

The output of this study the assessment of ABPRS as sampling frame that was used in both TDHS 2008 and TDHS 2013 and compilation of listing frame and the other frame which obtained from ABPRS.

<u>Points to consider for future surveys:</u>

i) Sampling of quarters in villages rather than blocks based on housing numbers that cannot be observed in the field would be better. If TURKSTAT provided us name of quarter instead of outdoor numbers, we would list all the quarter instead of searching for right numbers by spending time. This cause errors and physical effort as well. Of course this is under the assumption that the numbers of households in quarters are known.

ii) Another suggestion is the old school method as followed in TDHS-2008:.Either the whole village can be listed, or segmentation could be followed during the listing operation. Thus, no effort would be spent on searching for housing numbers.

<u>What to do for next survey?</u>

i) The existence of enumeration area can solve the big part of the problem, especially to create better frames.

ii) Numberings and name of streets should be updated; it will be helpful for listing operation definitely.

Finally, can the ABPRS be used as frame entirely? Once we look at the village instance, the answer is no. However, in provinces, if the numbers, which is given by TURKSTAT and listed by operations, are almost same; the answer turns into yes.