RE-CALCULATING THE SAMPLING VARIANCES OF SELECTED DEMOGRAPHIC INDICATORS FROM TDHS SURVEYS: THE EFFECT OF STRATIFICATION*

Tuğba ADALI **
A. Sinan TÜRKYILMAZ***

Turkey has a long tradition of demographic surveys. These surveys are essential in terms of producing official statistics regarding fertility, family planning and early age mortality. Moreover, they are complex sample surveys with multistage, stratified, clustered and weighted designs. The computation of sampling variances is not straight forward as for simple random samples for such complex samples. Clustering, weighting and stratification should be specified to the statistical software in order to obtain statistically correct sampling variances. This study is based on the method of calculation applied to calculate the sampling variances of rates in the Demographic and Health Surveys. We re-calculate the standard errors of the total fertility rate with a Jackknife n procedure that takes the stratification aspect into account and rather than the conventional DHS Jackknife 1 procedure and compare the findings. Our results show that the inclusion of stratification in the variance estimation process decreases the sampling variances/errors.

INTRODUCTION

Turkey has a long tradition of demographic surveys. National demographic surveys have been carried out by the Hacettepe University Institute of Population Studies (HUIPS) every five years since 1968. The last four surveys (through 1993 to 2008) were implemented as a part of the international Demographic and Health Surveys (DHS) project. These surveys are essential in terms of producing official statistics regarding fertility, family planning and early age mortality.

Turkey Demographic and Health Surveys are complex sample surveys with multistage, stratified, clustered and weighted designs. For such complex samples, the computation of sampling variances, and thus standard errors are not straight forward as for simple random samples. Clustering, weighting and stratification should be specified to the statistical software in order to obtain statistically correct sampling variances.

Many statistical software are currently capable of producing correct complex sample estimates (SAS, SPSS, Stata, etc.). The sampling errors for Turkey Demographic and Health Surveys (TDHS) have been calculated by Macro Int. by a special software designed specifically for DHS surveys: Integrated System for Survey Analysis (ISSA). ISSA has a module that calculates standard errors for means and proportions using Taylor series approximation, and has another module that calculates these statistics for demographic rates using Jackknife Repeated Replications (JRR).

---

* This paper is based on the PhD dissertation of the first author.

** Research Assistant, Hacettepe University, Institute of Population Studies
*** Assoc. Prof. Dr., Hacettepe University, Institute of Population Studies

While the main reports of TDHS surveys present the standard errors of means and proportion calculated by the Jackknife n method taking stratification, weighting and clustering into account, those of rates is calculated using Jackknife 1, which assumes there is a single stratum; as presented in the DHS sampling manuals (1996; 2012). This paper argues that JRR standard errors should be calculated using the same complex sample data information as Taylor series approximation does, and thus re-calculates the sampling variance and standard errors of the total fertility rate (TFR) from TDHS-1998, TDHS-2003 and TDHS-2008. Using the same surveys, the standard errors for a mean (ideal number of children) and a proportion (proportion of women with no education) are also calculated using the two different approaches of JRR to compare with the findings for TFR. The standard errors of the mean and proportion are also calculated with Taylor series approximation to cross check with the TDHS reports.

DATA

The Turkey Demographic and Health Surveys are similar in terms of sampling design. They all have weighted, multistage, stratified clustered sample designs. The sample sizes increased especially after TDHS-1998 due to renewed sampling needs. TDHS-1993 is not used in this paper, because the standard errors of TFR were not presented in its main report, therefore no comparisons can be made. The three surveys that are used are TDHS-1998, TDHS-2003 and TDHS-2008.

The samples of all three surveys were based on three stage selections in the urban areas. Settlements were selected from the frames shown in Table 1 at the first stage. At the 2nd stage, blocks of about 100 household each were selected by the Turkish Statistical Institute. Listing was done for all four surveys to update the blocks provided by Turkish Statistical Institute (TURKSTAT). A fixed take of 25 was selected from the lists at the last stage.

For most rural areas (rural areas without municipalities), TURKSTAT was not able to provide blocks, due to lack of lists. Lists were created from scratch for such settlements during listing. The fixed cluster size for rural areas was 15.

The fixed cluster take approach is one that is recommended by DHS (Macro International Inc., 1996), and adjusted for field operations in Turkey. The size of interviewing teams, the number individual interviews per household, the average length of interviews and time to travel to villages are were all taken into account to specify the fixed cluster takes in TDHS surveys.

Stratification is crucial in large sample surveys since it ensures a good spread of the sample across strata and improving representation. Geographical stratification was made for all four surveys. The strata for TDHS-1998 were based on 14 sub-regions of the five regions, doubled to differentiate urban and rural residences. The number of strata increased in the later surveys to maintain the 5 regions and provide estimates for the 12 regions as mentioned above. For TDHS-2003, there were even more sampling necessities: Special information was required for the earthquake regions, as well as the slum areas of Turkey. Thus the number of strata was the highest for this survey.

Table 1. Survey and Sample Information for TDHS-1998, TDHS-2003 and TDHS-2008

	TDHS-1998	TDHS-2003	TDHS-2008
Survey Implementation	HUIPS	HUIPS	HUIPS
Collaborating Organizations	General Directorate of Mother and Child Health and Family Planning , United Nations Population Fund (UNFPA) and Macro International	General Directorate of Mother and Child Health and Family Planning	T.R. Ministry of Health and T.R. Prime Ministry Undersecretary of State Planning Organization
Funding	United Nations Population Fund (UNFPA) and Macro International	State Planning Organization and the European Union.	Scientific and Technological Research Council of Turkey
Frame for first stage selection	1997 Population Count	2000 Census of Population	2007 Address Based Registration System
Estimation domains	5 regions, urban-rural	5 regions, 12 regions, urban-rural	5 regions, 12 regions, urban-rural
Questionnaires	HH, women aged 15-49, husbands	HH and EM Women aged 15-49	HH and EM Women aged 15-49
Number of strata	28	40	36
Number of clusters in the design	480	700	634
Number of completed HH interviews	8059	10,836	10,525
Number of completed interviews of EW	6512	8075	7405
HH response rate	93.80%	92.90%	88.40%
EW response rate	90.6% (all w)	95.60%	92.50%

Source: (Hacettepe Üniversitesi Nüfus Etütleri Enstitüsü, 1999; Hacettepe Üniversitesi Nüfus Etütleri Enstitüsü, 2005; Hacettepe Üniversitesi Nüfus Etütleri Enstitüsü, 2009)

Weighting in TDHS surveys serve two purposes: 1) To correct for disproportional allocation among strata to obtain unbiased statistics for the whole sample, 2) To correct for non-response in household and women interviews. The first component of the weight is obtained by the inverse of the selection probability, and the second component is obtained by response rates. Weights in TDHS surveys are strata level, meaning all units within the same stratum get the same sampling weight. The sum of weights is calibrated to match the size of completed interviews for both household and women's data sets.

All three surveys allow for the estimations for various domains: Urban and rural residences and five demographic regions. The definition of urban areas is very similar among surveys: Settlements with populations above 10,000 were considered urban[i]. The five regions have been conventionally used in demographic surveys since 1960s. They are named West, North, Central, South and East Anatolia. A slight change has been made in their geographical definitions from 1998 to 2003. This change was brought about in 2002, in accordance with the accession process of Turkey to the European Union. According to the Nomenclature of Territorial Units for Statistics (NUTS) system, a new three level region classification was prepared for Turkey. The NUTS 1 level consists of 12 regions, and official statistics are now produced for these regions as well. The conventional borders of the 5 regions were altered so that it would be possible to provide estimates for both these classical regions, and the new 12 regions.

METHODS

As mentioned earlier, the total fertility rate re-calculated in this study is more complex than its formal definitions found in textbooks; due to the introduction of exposures of women. Its DHS definition will be provided before proceeding with the adjustment made in the sampling variance estimation process. Two other indicators are used in the results section to compare the findings obtained for the TFR: Ideal number of children and the proportion of women with no education. These indicators were defined as they were defined in the main TDHS reports. The findings for these are obtained using the SAS procedures PROC SURVEYMEANS and PROC SURVEYFREQ which allow complex sample estimations. Type of variance estimation is specified by the VARMETHOD option.

Total Fertility Rate

Total fertility rates in TDHS surveys are calculated from exact dates collected in birth histories of women. Conventionally, the denominators of age specific fertility rates (the sum of which make up the total fertility rate) are defined as the midyear population of women in the corresponding age groups. In DHS surveys, the denominator is the sum of person years spent in the age group by all women in the sample.

The total fertility rate is not really a rate, but rather a sum of rates. It consists of the summation of age specific fertility rates (ASFR), usually calculated for seven different 5 year age groups, covering a span of 35 years in total. These are then summed up to make up the total fertility rate, which can be interpreted as the expected number of live births a woman would give through her reproductive ages provided that she experiences the current age specific fertility rates. Thus the TFR, calculated from a synthetic cohort is a period measure.

The TFR is calculated as:

$$TFR = \sum_{i=1}^{7} ASFR(i,t)$$

The age specific rates are calculated as follows:

$$ASFR(i,t) = \frac{b(i,t)}{e(i,t)} = \frac{\sum_j w_j I_{[a,b]}}{\sum_j awf_j EXP_{[a,b]}}$$

where i denotes age group, t denotes the length of the specific period on which the calculation is based, b(i,t) denotes births to women in age group i during t, and e(i,t) denotes the exposure of women in age group i in person years during time t where $I_{[a,b]}$ is an indicator that has the values:

$$I_{[a,b]} = \begin{cases} 1 \text{ if a} < \text{mother's age at birth} < \text{b and } < \text{child's age at interview date} < t; \\ 0 \text{ Otherwise} \end{cases}$$

For the denominator:

$$EXP_{[a,b]} = \begin{cases} [\text{b} - \max(\text{a}, \text{age}_{\min})] & \text{if age}_{\min} < \text{b} < \text{age}_{\max} \\ [\text{age}_{\max} - \max(\text{a}, \text{age}_{\min})] & \text{if a} < \text{age}_{\max} < \text{b} \end{cases} \quad \text{(Thiam ve Aliaga, 2001).}$$

Jackknife Repeated Replications

Whenever cluster sizes depart from being equal sized, and/or there is a departure from equal probability selection (epsem; i.e. a self-weighted design), weighting is introduced. While statistics are easily calculated from simple random samples, they become weighted quantities in such complex samples. Any mean, proportion or rate, is no longer a simple mean, but rather a *ratio mean*, where the denominator (sample size) is a random variable, because of varying weights.

Variance estimation for complex sample statistics is therefore not straightforward. There are several ways of tackling this issue. In DHS surveys, Taylor Linearization Method is widely used for the calculation of variances for means and proportions, provided they are faster in computation than replication methods (ICF International, 2012). For demographic rates, however, replication methods are more suitable, since they are more complicated statistics.

The replication method used in DHS surveys is the Jackknife Repeated Replications Method. The idea in this method is to calculate the overall statistic, and then remove and replace units one by one, re-calculate the overall statistic from each subsample, and calculate the overall variance of the statistics produced by each subsample.

Although based on the same formula, there are often three different procedures recognized for implementing the Jackknife Repeated Replications (JRR) method, depending on stratification. If there is no stratification (i.e. there is only one stratum), the JRR employed is called Jackknife 1 (Westat, 2007). In this case, as each cluster is deleted, the rest will be reweighted so that the sum of weights will remain the same as original. If the survey has a paired sample design, then the JRR employed is called Jackknife 2. Here, once a cluster is dropped, the other cluster in the same stratum will be doubled in weight. The last procedure is called Jackknife n, which allows the number of clusters in a stratum to be 2 or more. Whenever a cluster is deleted, the rest of the clusters in the stratum are weighted up to compensate.

The JRR used by DHS for Turkey DHSs to calculate the variance of rate r is Jackknife 1, given below:

$$SE^2(r) = var(r) = \frac{1}{k(k-1)}\sum_{i=1}^{k}(r_i - r)^2 \qquad (1)$$

in which

$$r_i = kr - (k-1)r_{(i)}$$

and

k is the total number of clusters,

r is the estimate (in this case, the rate) computed from the full sample of k clusters, and

r(i) is the replicate estimate computed from the replicate sample of k-1 clusters (i^{th} cluster is excluded) (Aliaga ve Türkyilmaz, 1999; ICF International, 2012; Türkyılmaz, 2004; Türkyılmaz ve Adalı, 2009).

This formula drops one cluster at a time, thus forms as many replicates as the number of clusters; excluding any data on stratification.

The way DHS deals with stratification for the sampling variances of proportions and rates is through implicit stratification. This means the variable supplied to software for stratification is not one that signifies the design strata. In Turkish DHSs, there is geographical ordering of settlements within design strata at the stage of systematic selection (probability proportional to size – PPS). This ensures a well spread among different geographical locations within the design strata, creating implicit strata; the PSUs in which have geographical proximity. DHS uses this property, and creates strata consisting of two PSUs each, like a paired selection design, and uses these strata when performing Taylor Linearization for variance estimation. First of all, these paired strata were generated to cross check with DHS, and then they were used for JRR, too. Each design strata has an even number of PSUs for this purpose in TDHS surveys, except for instances of PSU level non-response, where this number can become an odd number. Jackknife Repeated Replications can function with either. Whenever a PSU is dropped from a stratum, the rest of the units are weighted up in a way that they still have as large a sampling weight in total as the original stratum has.

As Yıldız (2011) stated in her work regarding standard errors in TDHS surveys, the Jackknife n method is the most suitable approach for TDHS surveys; contrary to the DHS use of Jackknife 1. Westat (2007) suggests that Jackknife be used when explicit stratification has not been used; which has actually been used in every TDHS so far. Moreover, DHS documentations lack an explanation on why stratification is accounted for with Taylor Series approximation, but not JRR.

The edited JRR formula (Jackknife n) for the demographic rates in TDHS surveys is as follows (Lee ve Forthofer, 2006):

$$SE^2(r) = var(r) = \sum_{\alpha=1}^{a} \frac{b(_\alpha - 1)}{b_\alpha} \sum_{\beta=1}^{b_\alpha} (r_{\alpha\beta} - r)^2 \tag{2}$$

where

r is the overall rate without any deletion,

$r_{\alpha\beta}$ is the rate computed on the replicate from strata α, where cluster β is deleted, and the remaining ($b_{\alpha-1}$) elements are weighted up by the factor $b_\alpha/(b_\alpha - 1)$. The results provided in the next section compare the findings obtained with formulas 1 and 2.

The results section compares the findings for TFR for Turkey and the five main demographic regions of the country. The calculations were made using SAS, by calling a SAS macro. This SAS macro was previously written by Thiam and Aliaga (2001) for the sake of sampling variance estimation. It calculates the sampling variances with Jackknife 1; this Macro was adjusted to calculate it with Jackknife n, along with other adjustments not related to this paper.

The complex sample standard errors of the two variables used to cross check the findings of TFR (a mean and a proportion) are calculated by both using a JRR 1 approach, and a JRR n approach. This was ensured by creating a variable with a constant value for each observation, and defining it as the stratum variable for JRR 1. The implicit strata used by DHS for Taylor Series Estimations is used as the stratum variable for JRR n. For the mean and proportion; the results from Taylor series approximation[ii] are also presented, because they are the default in DHS surveys. Both the Taylor series standard errors with one stratum and multiple strata are presented to support the phenomenon observed with JRR.

RESULTS

The point estimates for all tables presented here are the same as those presented in the TDHS reports. This consistency shows that the SAS Macro and SAS procedures function as they should. The standard errors of the 3 year TFRs calculated by the SAS Macro (JRR n) are lower than those given in the report. However, these differences are small. This leads to somehow narrower confidence intervals.

Table 2. Estimates of 3 year TFRs by region, TDHS-1998

		Estimate	SE	CV	Confidence Interval Lower	Upper	% Change in SE
Turkey	**JRR 1 (DHS)**	**2.609**	**0.079**	**0.030**	**2.451**	**2.768**	**-0.10**
	JRR n (SAS)	*2.609*	*0.071*	*0.027*	*2.468*	*2.751*	
West	**JRR 1 (DHS)**	**2.032**	**0.122**	**0.060**	**1.788**	**2.277**	**-0.04**
	JRR n (SAS)	*2.032*	*0.117*	*0.058*	*1.799*	*2.266*	
South	**JRR 1 (DHS)**	**2.555**	**0.15**	**0.059**	**2.254**	**2.864**	**-0.04**
	JRR n (SAS)	*2.555*	*0.144*	*0.056*	*2.267*	*2.843*	
Central	**JRR 1 (DHS)**	**2.564**	**0.132**	**0.051**	**2.299**	**2.828**	**0.03**
	JRR n (SAS)	*2.564*	*0.136*	*0.053*	*2.292*	*2.835*	
North	**JRR 1 (DHS)**	**2.679**	**0.173**	**0.065**	**2.333**	**3.025**	**-0.05**
	JRR n (SAS)	*2.679*	*0.165*	*0.062*	*2.35*	*3.008*	
East	**JRR 1 (DHS)**	**4.191**	**0.225**	**0.054**	**3.741**	**4.641**	**-0.08**
	JRR n (SAS)	*4.191*	*0.208*	*0.050*	*3.776*	*4.607*	

The findings for TDHS-2003 are similar to those of TDHS-1998 (Table 3). The point estimates are the same, but the standard errors are slightly lower in almost all regions. The resulting confidence intervals are again slightly narrower. The regional absolute changes are slightly lower than the overall absolute change in both surveys.

Table 3. Estimates of 3 year TFRs by region, TDHS-2003

	JRR type	Estimate	SE	CV	Confidence Interval Lower	Upper	% Change in SE
Turkey	**JRR 1 (DHS)**	**2.231**	**0.054**	**0.024**	**2.123**	**2.339**	**-0.11**
	JRR n (SAS)	*2.231*	*0.048*	*0.022*	*2.135*	*2.327*	
West	**JRR 1 (DHS)**	**1.879**	**0.068**	**0.036**	**1.744**	**2.015**	**-0.10**
	JRR n (SAS)	*1.879*	*0.061*	*0.032*	*1.757*	*2.002*	
South	**JRR 1 (DHS)**	**2.297**	**0.159**	**0.069**	**1.979**	**2.616**	**-0.04**
	JRR n (SAS)	*2.297*	*0.152*	*0.066*	*1.994*	*2.601*	
Central	**JRR 1 (DHS)**	**1.864**	**0.098**	**0.053**	**1.667**	**2.060**	**-0.04**
	JRR n (SAS)	*1.864*	*0.094*	*0.050*	*1.676*	*2.052*	
North	**JRR 1 (DHS)**	**1.942**	**0.119**	**0.061**	**1.704**	**2.180**	**0.06**
	JRR n (SAS)	*1.942*	*0.126*	*0.065*	*1.690*	*2.194*	
East	**JRR 1 (DHS)**	**3.645**	**0.162**	**0.044**	**3.331**	**3.978**	**-0.03**
	JRR n (SAS)	*3.654*	*0.157*	*0.043*	*3.340*	*3.969*	

For TDHS-2008, the overall standard error for Turkey is very close under two variance estimation methods (Table 4). The largest regional difference is observed as -0.9%. Two of the regional changes were positive, yet very small. Unlike the other two surveys, the overall absolute difference in standard errors is smaller than those in regional standard errors.

Table 4. Estimates of 3 year TFRs by region, TDHS-2008

	JRR type	Estimate	SE	CV	Confidence Interval Lower	Upper	% Change in SE
Turkey	**JRR 1 (DHS)**	**2.156**	**0.059**	**0.027**	**2.038**	**2.273**	**-0.02**
	JRR n (SAS)	*2.156*	*0.058*	*0.027*	*2.040*	*2.272*	
West	**JRR 1 (DHS)**	**1.734**	**0.099**	**0.057**	**1.536**	**1.932**	**0.05**
	JRR n (SAS)	*1.734*	*0.104*	*0.060*	*1.526*	*1.941*	
South	**JRR 1 (DHS)**	**2.092**	**0.115**	**0.055**	**1.862**	**2.322**	**-0.06**
	JRR n (SAS)	*2.092*	*0.108*	*0.052*	*1.875*	*2.309*	
Central	**JRR 1 (DHS)**	**2.198**	**0.122**	**0.056**	**1.955**	**2.442**	**-0.09**
	JRR n (SAS)	*2.198*	*0.111*	*0.051*	*1.976*	*2.421*	
North	**JRR 1 (DHS)**	**2.082**	**0.117**	**0.056**	**1.848**	**2.316**	**0.04**
	JRR n (SAS)	*2.082*	*0.122*	*0.058*	*1.839*	*2.326*	
East	**JRR 1 (DHS)**	**3.274**	**0.142**	**0.043**	**2.990**	**3.557**	**-0.07**
	JRR n (SAS)	*3.274*	*0.133*	*0.040*	*3.009*	*3.539*	

The ideal number of children ever born and its standard error are estimated using SAS for all three surveys to simulate the effect of stratification in JRR variance estimation of means; by defining a single stratum and multiple strata respectively (Table 5). The same is applied for the Taylor series variance, and virtually the same results are found.

The results are consistent with the findings for TFR in terms of the direction of change in standard errors. Taking stratification into account decreases these statistics. However, the magnitude is different; the decrease in standard errors observed for this variable is greater. For TDHS-2003 and TDHS-2008 the decrease is almost by a quarter.

Table 5. Estimates of ideal number of children ever born per woman, TDHS-1998*, TDHS-2003, TDHS-2008

	Type of variance estimation	Estimate	SE	CV	Confidence Interval Lower	Upper	% Change in SE
1998	JRR 1	2.36	0.022	0.009	2.31	2.40	**-0.22**
	JRR n	**2.36**	**0.017**	**0.007**	**2.32**	**2.39**	
	Taylor 1	*2.36*	*0.022*	*0.009*	*2.31*	*2.40*	
	Taylor n (DHS)	*2.36*	*0.017*	*0.007*	*2.32*	*2.39*	
2003	JRR 1	2.51	0.025	0.010	2.46	2.56	**-0.19**
	JRR n	**2.51**	**0.020**	**0.008**	**2.47**	**2.55**	
	Taylor 1	*2.51*	*0.025*	*0.010*	*2.46*	*2.56*	
	Taylor n (DHS)	*2.51*	*0.020*	*0.008*	*2.47*	*2.55*	
2008	JRR 1	2.52	0.025	0.010	2.47	2.57	**-0.15**
	JRR n	**2.52**	**0.021**	**0.008**	**2.48**	**2.56**	
	Taylor 1	*2.52*	*0.025*	*0.010*	*2.47*	*2.57*	
	Taylor n (DHS)	*2.52*	*0.021*	*0.008*	*2.48*	*2.56*	

** The denominators for TDHS-1998 include all women, the rest are based on ever married women for compliance with TDHS main report.*

The proportion of women with no education and its standard error are estimated using SAS as an example of proportions (Table 6). These results are also consistent with the findings for rates and means, observed in the previous tables; portraying decreasing standard errors as the method changes from JRR 1 to JRR n. The same values are observed for the Taylor series standard errors. The magnitude of decrease is closer to what is observed for the ideal number of children, i.e. the change is larger.

DISCUSSION

This paper was based on the fact that the complex sample variance estimation technique for demographic rates in DHS surveys, the Jackknife 1 procedure, assumes a single stratum. It aimed to investigate what effect this sampling property has on the calculated standard errors of one of the most widely used demographic rate: the total fertility rate. The JRR and Taylor series standard errors of two simpler statistics (a mean and a proportion) were also included to cross check the results.

A SAS Macro previously written by Thiam and Aliaga (2001) was obtained to calculate the sampling variance of the TFR. The Jackknife Repeated Replications method included in this macro was a Jackknife 1 type, one that is based on clustering, and not stratification. This macro was first run to make sure the estimates are the same as those presented in TDHS reports, and were then modified to take stratification into account (Jackknife n).

Table 6. Estimates of proportion of women with no education*, TDHS-1998, TDHS-2003, TDHS-2008**

	Type of variance estimation	Estimate	SE	CV	Confidence Interval Lower	Confidence Interval Upper	% Change in SE
1998	JRR 1	16.74	0.871	0.05	15.029	18.444	**-0.31**
	JRR n	**16.74**	**0.604**	**0.04**	**15.553**	**17.921**	
	Taylor 1	*16.74*	*0.871*	*0.05*	*15.031*	*18.443*	
	Taylor n (DHS)	*16.74*	*0.604*	*0.04*	*15.553*	*17.921*	
2003	JRR 1	21.81	1.070	0.05	19.714	23.908	**-0.24**
	JRR n	**21.81**	**0.817**	**0.04**	**20.211**	**23.412**	
	Taylor 1	*21.81*	*1.068*	*0.05*	*19.718*	*23.904*	
	Taylor n (DHS)	*21.81*	*0.817*	*0.04*	*20.211*	*23.412*	
2008	JRR 1	18.34	1.006	0.05	16.367	20.310	**-0.20**
	JRR n	**18.34**	**0.808**	**0.04**	**16.756**	**19.922**	
	Taylor 1	*18.34*	*1.004*	*0.05*	*16.371*	*20.307*	
	Taylor n (DHS)	*18.34*	*0.800*	*0.04*	*16.770*	*19.908*	

** The definition for "no education" includes those who have not attended school at all in TDHS-1998, and those who have not completed the first 5 years of compulsory education in the later surveys for compliance with the TDHS main reports.*
***The denominators for TDHS-1998 include all women, the rest are based on ever married women for compliance with the TDHS main reports.*

Results for the TFR calculations showed that the point estimates provided by DHS and the SAS macros are identical, provided the method of calculation is the same. The standard errors computed by the Jackknife n method were generally lower than those by Jackknife 1. The percent decrease was 11% at most (Table 3). The results for the mean (ideal number of children for interviewed women) and the proportion (proportion of interviewed women with no education) were parallel to those for TFR in the sense that the JRR n standard errors were lower than JRR 1 standard errors. The magnitude of decrease, on the other hand, was higher for the mean and proportion, reaching up to almost a quarter at times. These differences are related to the natures of these statistics (e.g. the rate of homogeneity). While one is a sum of rates (TFR), the other is a mean (ideal number of children); and the last one is a proportion (proportion of women with no education).

The reason for the decrease in the sampling variance when stratification is included is no different than why stratification ensures lower sampling variances compared to simple random sampling. Omitting stratification should overestimate the sampling variances in theory. Using JRR 1 rather than JRR n in DHSs means using stratification for design reasons – a better spread thus strong representation of the population – and not taking advantage of the fact that it reduces sampling error. The implicit strata in TDHS are homogenous within, both because they are in the same explicit design strata, and because they are paired/tripled from an ordered PSU list. In any given strata, the settlements are sorted and then grouped into implicit strata, ensuring the PSUs within are homogenous within. Instead of dropping one unit, and replacing it with all remaining units (as in JK 1), any dropped unit is replaced with a similar unit in JK n, when stratification is considered.

Although the standard errors were not very different and a computational burden is introduced, we suggest that stratification be included in calculations for the sake of methodology. The JK 1 method is only suitable when there is no stratification involved (Westat, 2007); which is contradictive to the general sampling design of demographic sample surveys in Turkey. This study agrees with Yıldız (2011) that JK n is the most suitable for TDHS surveys. Besides, no reason is provided in the DHS manuals (1994; 2012) as to why stratification is included in the Taylor series variance estimations but not Jackknife variance estimations. However, using a JK 1 approach provides more liberal confidence intervals, which make the results less assertive; making room for any possible non-sampling error that cannot be foreseen.

Acknowledgements

We are grateful to Mamadou Thiam for providing us with the SAS Macro used by Thiam and Aliaga (2001). We also thank the Hacettepe University Institute of Population Studies for the TDHS data sets.

REFERENCES

Aliaga, A., Türkyilmaz, A.S. (1999). "Appendix C Estimates of Sampling Errors". **1998 Turkey Demographic and Health Survey**. Ankara, Turkey: Hacetepe University.

Hacettepe Üniversitesi Nüfus Etütleri Enstitüsü. (1999). **Türkiye Nüfus ve Sağlık Araştırması 1998**. Ankara: Hacettepe Üniversitesi Nüfus Etütleri Enstitüsü, Macro International Inc., Sağlık Bakanlığı Ana Çocuk Sağlığı ve Aile Planlaması Genel Müdürlüğü, Birleşmiş Milletler Nüfus Fonu, Amerika Birleşik Devletleri, Uluslararası Kalkınma Teşkilatı.

Hacettepe Üniversitesi Nüfus Etütleri Enstitüsü. (2005). **Türkiye Nüfus ve Sağlık Araştırması 2003**. Ankara: Hacettepe Üniversitesi Nüfus Etütleri Enstitüsü, Sağlık Bakanlığı Ana Çocuk Sağlığı ve Aile Planlaması Genel Müdürlüğü, T.C. Başbakanlık Devlet Planlama Teşkilatı, Avrupa Birliği.

Hacettepe Üniversitesi Nüfus Etütleri Enstitüsü. (2009). **Türkiye Nüfus ve Sağlık Araştırması 2008**. Ankara: Hacettepe Üniversitesi Nüfus Etütleri Enstitüsü, Sağlık Bakanlığı Ana Çocuk Sağlığı ve Aile Planlaması Genel Müdürlüğü, Başbakanlık Devlet Planlama Teşkilatı.

ICF International. (2012). **Demographic and Health Survey Sampling and Household Listing Manual**. Calverton, Maryland USA: Measure DHS.

Lee, E.S., Forthofer, R.N. (2006). "Strategies for Variance Estimation". **Analyzing Complex Survey Data**: SAGE.

Macro International Inc. . (1996). **Sampling Manual**. Calverton, Maryland, USA: Macro International Inc.

Thiam, M., Aliaga, A. (2001). **Estimating Sampling Errors of Means, Total Fertility, and Childhood Mortality Rates Using SAS**. ORC Macro.

Türkyılmaz, A.S. (2004). "Appendix C Estimates of Sampling Errors". **2003 Turkey Demographic and Health Survey**. Ankara, Turkey: Hacettepe University.

Türkyılmaz, A.S., Adalı, T. (2009). **Appendix C Sampling Errors**. Ankara: Hacettepe University Institute of Population Studies, T.R. Ministry of Health Directorate of Mother and Child Health, T.R. Prime Ministry State Planning Organizaiton.

WesVAr. (2007). **WesVar 4.3 User Guide**.

Yıldız, D. (2011). **Sampling Error Estimation by Using Different Methods and Software in Complex Samples**. Ankara: Hacettepe University.

Notes

[i] The exception is TDHS-1993, where district centers-regardless of population size-were considered urban due to their administrative status.
[ii] See (Macro International, 1996; ICF International, 2012) for information on Taylor series approximation for variance estimation.

ÖZET

TNSA VERİLERİNDEN SEÇİLMİŞ DEMOGRAFİK GÖSTERGELERİN TEKRAR HESAPLANMASI: TABAKALAMANIN ETKİSİ

Türkiye'de köklü bir demografik araştırma geleneği bulunmaktadır. Bu araştırmalar doğurganlık, aile planlaması ve erken yaş ölümlülüğüyle ilgili resmi istatistik üretmek için önem taşımaktadır. Çok aşamalı, tabakalı, kümeli ve ağırlıklı örneklem yaklaşımlarıyla tasarlamış bu araştırmalar, karmaşık örneklem (complex sample) tasarımlı araştırmalar olmaları nedeniyle örneklem varyansları özel yaklaşımlarla hesaplanmaktadır. Karmaşık örneklemlerde bu hesaplamalar basit tesadüfi seçilmiş örneklemlerdeki gibi hesaplanamamaktadır. Kullanılan istatistik yazılımına örneklem varyanslarının doğru hesaplanabilmesi için tabaka, küme ve ağırlık bilgileri tanıtılmalıdır. Bu çalışma Nüfus ve Sağlık Araştırmalarında uygulanan örnekleme varyansı hesabından yola çıkmaktadır. Geleneksel olarak Jackknife 1 yöntemiyle hesaplananan toplam doğurganlık hızının standard hatasını Jackknife n yöntemiyle hesaplayıp, bulguların karşılaştırılması amaçlanmıştır. Sonuçlar tabakalamanın hesaba katılmasının örnekleme varyansını/hatalarını düşürdüğünü göstermektedir.