HACETTEPE UNIVERSITY INSTITUTE OF POPULATION STUDIES
DEPARTMENT OF DEMOGRAPHY

POOLING TURKEY DEMOGRAPHIC AND HEALTH SURVEY DATA INTO AN AGGREGATED DATA SOURCE: IMPLICATIONS ON SELECTED DEMOGRAPHIC INDICATORS

Tuğba ADALI

Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of Doctor of Philosophy in
Demography at Hacettepe University
Institute of Population Studies

Ankara
March, 2014

HACETTEPE UNIVERSITY INSTITUTE OF POPULATION STUDIES
DEPARTMENT OF DEMOGRAPHY

POOLING TURKEY DEMOGRAPHIC AND HEALTH SURVEY DATA INTO AN AGGREGATED DATA SOURCE: IMPLICATIONS ON SELECTED DEMOGRAPHIC INDICATORS

Tuğba ADALI

Supervisors
Prof. Dr. James M. Lepkowski
Assoc. Prof. Dr. Ahmet Sinan Türkyılmaz

Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of Doctor of Philosophy in
Demography at Hacettepe University
Institute of Population Studies

Ankara
March, 2014

ACCEPTANCE AND APPROVAL

This is to certify that we have read and examined this thesis and that in our opinion it is fully adequate, in scope and quality as a thesis for the degree of Doctor of Philosophy in Demography.

<u>Examining Committee Members:</u>

Member (Chair): ....
Assoc. Prof. Dr. Özlem İlk Dağ, Middle East Technical University, Department of Statistics

Member (Supervisor): ....
Assoc. Prof. Dr. Ahmet Sinan Türkyılmaz, Hacettepe University, Institute of Population Studies

Member: ....
Prof. Dr. Banu Akadlı Ergöçmen, Hacettepe University, Institute of Population Studies

Member: ....
Prof. Dr. İsmet Koç, Hacettepe University, Institute of Population Studies

Member: ....
Assoc. Prof. Dr. Erdem Karabulut, Hacettepe University, Department of Biostatistics

This thesis has been accepted by the above-signed members of the Jury and has been confirmed by the Administrative Board of the Institute of Population Studies, Hacettepe University.

Date: .... / .... / 2014

....
Prof. Dr. Ş. Armağan Tarım
Director

to beloved Hande, a
demographer lost too
soon…

ABSTRACT

This dissertation aimed to pool data from Turkey Demographic and Health Surveys with the aim to improve the precision of various demographic indicators. Increased precision was expected to prove useful, especially for region level estimates and rate estimates based on shorter time periods.

Turkey has a rich supply of demographic survey data obtained from the surveys carried out by the Hacettepe University Institute of Population Studies since the 1960s. These surveys provided Turkey with a tradition of sample surveys as well as enabling the accumulation of data which is seldom found in other countries. Surveys are among the major sources of demographic indicators in Turkey. Four surveys conducted within the international Demographic and Health Surveys Program were used in this dissertation.

The first phase of the thesis consisted of studying the methodological issues regarding the pooling of multiple complex sample datasets. These issues included sampling weight adjustments, and complex sample standard error estimation of rates. Two proportions, two means and two rates were selected to find out the implications of the aggregated data set on these indicators.

Pooling survey data was found to improve precision for demographic indicators, decreasing standard errors for Turkey as a whole and within regions. Results generally provided stable point estimates throughout different combinations of surveys, and confidence intervals narrowed down demonstrating higher precision. It was concluded that this procedure could be useful so long as specific attention was paid to certain data quality and bias related to target populations.

ÖZET

Bu tez Türkiye Nüfus ve Sağlık Araştırmaları verilerini birleştirerek çeşitli demografik göstergeler için istatistiksel kesinliği iyileştirmeyi amaçlamıştır. Kesinliğin iyileşmesinin özellikle bölge düzeyi tahminler ve daha kısa dönemlerden hesaplanan hızlar için işlevli olması beklenmiştir.

Türkiye 1960'lardan beri Hacettepe Üniversitesi Nüfus Etütleri Enstitüsü tarafından yürütülmekte araştırmalar sayesinde demografik araştırmalar açısından zengin bir ülkedir. Bu araştırmalar Türkiye'de yerleşmiş bir örneklem araştırması geleneğini oluştururken, benzeri başka ülkelerde sıklıkla görülmeyen bir veri birikiminin oluşmasını da sağlamıştır. Araştırmalar Türkiye'de demografik göstergelerin temel kaynakları arasındadır. Bu tezde uluslararası Demographic and Health Surveys Program (Nüfus ve Sağlık Araştırmaları Programı) çerçevesinde gerçekleştirilmiş dört araştırmanın verisinden yararlanılmıştır.

Tezin ilk aşaması karmaşık örneklemli örneklem araştırmalarından veri birleştirmekle ilgili yöntemsel çalışmaları içermiştir. Bu çalışmalar örneklem ağırlığı düzeltmesi ve demografik hızların karmaşık örneklemlerde standart hatalarının hesaplanmasına ilişkindir. Birleştirilmiş verinin göstergelerdeki etkisini gözlemleyebilmek için iki oran, iki ortalama ve iki hız seçilmiştir.

Araştırma verilerinin birleştirilmesi sonucunda demografik göstergeler için hem Türkiye düzeyinde, hem de bölgeler düzeyinde istatistiksel kesinliğin iyileştiği görülmüştür. Sonuçlar genel olarak farklı veri setlerinin birleşimlerinden istikrarlı nokta tahminleri sağlamıştır. Bu tahminlerin güven aralıkları daralmıştır, bu da daha iyi kesinliği işaret etmektedir. Veri kalitesi ve hedef nüfuslarla ilgili olası yanlılıklar göz önünde tutulduğu sürece veri birleştirme prosedürünün yararlı olduğu sonucuna varılmıştır.

ACKNOWLEDGEMENTS

Having started my PhD as a 23 year old, and having defended my dissertation one day before my 30th birthday (technically still 29, demography rules!), I cannot possibly list all those I encountered during my journey. To keep it short:

I am grateful to my two advisors, Ahmet Sinan Türkyılmaz and James M. Lepkowski for being supportive and understanding. It is the greatest relief for a PhD student to know that she will have something new to think about whenever she asks for advice.

I am thankful to my thesis supervising committee members who encouraged me at each regular meeting; and the thesis defense jury members who provided me with feedback and corrections: Prof. Dr. Banu Akadlı Ergöçmen, Prof. Dr. İsmet Koç, Assoc. Prof. Dr. Ahmet Sinan Türkyılmaz, Assoc. Prof. Dr. Özlem İlk Dağ and Assoc. Prof. Dr. Erdem Karabulut.

I thank my dearest high school buddies (my two BFFs, M&Ms in particular - love you guys), my pomodoro buddy Gülnur (looking so good on Skype), my funny and loving family (with the cutest baby boy ever), my warm and smart friends and colleagues from HUIPS (Berna and Dilek: this is where you go), the exes that had to share the suffering during dissertation writing, the students at UoM / ISR / SMP back when I was in Michigan, and *internet* for my daily dose of procrastination: You never let me down.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION AND BACKGROUND

The purpose of this thesis is to combine the datasets of the Turkey Demographic and Health Surveys to create an alternative demographic data source and thus contribute to the enrichment of demographic analysis in Turkey.

Demographic data is obtained from censuses, vital registration systems, and sample surveys. While all three sources are present in Turkey, their characteristics mostly limit them to surveys when it comes to key fertility and mortality indicators. Having begun in the early years of the Republic of Turkey, censuses have been quinquennial until 1990. The last census, which appears to be the final one has been carried out in 2000. Censuses have been extremely useful in estimating population totals, age structure, spatial distribution and basic characteristics of the population, with some criticism regarding the overestimation of the population. Yet given their nature of cost and time, they were based on short questionnaires, like most of their foreign counterparts. The brief questionnaire allows little detail, thus widely used demographic indicators like the total fertility rate and infant mortality rate are often calculated indirectly from such data.

A form of vital registration system has been in use in Turkey since the late Ottoman era that would keep records of individuals born to each family (Shaw, 1978). These records later became digitalized to constitute the Central Civil Registration System (Merkezi Nüfus İdaresi Sistemi, MERNIS) (NVİ, 2014). However, this system was rather limited in terms of coverage, detail and being up-to-date. Moreover, it would not link persons to their current residences until the relatively new Address Based Population Registration System (Adrese Dayalı Nüfus Kayıt Sistemi, ADNKS) was launched in 2007. Since this date, basic characteristics of the de jure population have been compiled from this registration system and published every year. Yet the timeliness of the reporting of vital events is a potential issue for this system. For example, late reporting is documented for births from the birth history section in Turkey Demographic and Health Survey 2008 (HUIPS, 2009) .

Cross-sectional demographic surveys thus fill out an important gap in Turkey in terms of demographic data. They have been carried out since 1968 on a quinquennial basis, with extensive questionnaires on households and women of reproductive ages. They provide detailed information on fertility, reproductive health and infant mortality, allowing the direct estimation of fertility and mortality rates. These sample surveys have been in line with international trends of surveys: Knowledge, Attitude and Practice Surveys in the 1960s, World Fertility Surveys starting in the 1970s, Contraceptive Prevalence Surveys in the 1980s and Demographic and Health Surveys starting in the 1990s. The Demographic and Health Surveys in Turkey were carried out in 1993, 1998, 2003 and 2008. These surveys had major similarities in terms of their sampling designs and questionnaire contents. They also included mostly standard birth history sections which provide detailed records of women’s live birth experiences since their beginning of childbearing.

There are currently four DHS data sets available for Turkey. The main advantage expected from the cumulation process in this dissertation is to provide higher precisions, thus narrower confidence intervals for widely used demographic estimates. Different types of statistics are chosen to exemplify demographic indicators: Proportions, means and rates. The main argument is that with the improvement of statistical precision via larger data sets, estimations for smaller domains or shorter time intervals that normally would be avoided would be possible.

Combining, used interchangeably with cumulating or pooling surveys have become a popular topic of interest in the academic field especially in the last two decades, with the higher availability of data and standard survey series, according to reviewed literature. While examples of combining over space are common for international surveys such as the Demographic and Health Surveys; examples of combining across time are usually common for national surveys, such as the US National Health Interview Survey. Other examples of combining surveys include those between entirely different surveys to exploit a wide range of variables; between administrative records and surveys and more. The application in this dissertation falls into the category of combining national data across time.

Sampling errors in sample surveys are directly related to sample size. The larger the sample size, the smaller the standard errors for any estimate, under the same survey design. The sample sizes of the Demographic and Health Surveys are determined in a complicated manner. In addition to specific output needs of the country, there are recommendations by DHS for minimum sample sizes for certain indicators to follow. Oversampling is made for domains where higher precision is desired; there are budget constraints and logistic matters. In Turkey as a whole, estimates are based on 6000-8000 women for all four DHS surveys. However, total fertility rates tend to be calculated for 3 years, and infant mortality rates for 5 or 10 years depending on the domain of estimation, to keep the number of events sufficient for the rates to be stable. Birth history sections of questionnaires allow overlapping periods, allowing estimates to be calculated for the same time periods from different surveys. This increases the sample size, providing both a temporal, and a spatial advantage in terms of precision of estimates.

Demographic and Health Surveys have been an important source for academic research. Until recently, analyses from these data have been either in a cross-sectional sense, or an event-history approach by the utilization of the birth history section that supplies a flow of data. Separate datasets have not been analyzed at once until recently, the few examples of which will be mentioned in the literature review section. Combination of DHS data is usually performed across space rather than time, and applying the latter will allow more precise estimates for smaller domains and allow rates for shorter time intervals. Kish (1999) has suggested that the cumulation of periodic surveys must receive serious technical consideration in the future, claiming such work done have so far been by ad hoc procedures. Especially for cumulations of long periods such as 10 years, secular trends in population size must be considered (Kish, 1999). This thesis contributes to the literature as an example of combining the Turkey Demographic and Health Surveys using their complex sample properties, including weighting, at the same time.

Given the complex sample structures of the Turkey Demographic and Health Surveys, and the complexity of the demographic indicators themselves, this aim presents certain challenges. First of all, a standardization of the datasets is required. Secondly, given they are conventionally tabulated at Measure DHS; the sampling error computation of demographic rates under complex sample settings need to be figured out. Finally, new weights, stratification and clustering should be re-considered for the combined data set.

The dissertation proceeds with a literature review section. The literature review (Section 2) first of all covers the development of sampling theory and sample surveys in demography. This section continues with literature on combining surveys: Studies related to the conceptual issues regarding combining surveys and empirical examples are presented. Sample weight issues in these studies are presented separately. Section 3, the data and methods section first includes detailed information on the surveys that are used in this thesis. A separate section is on weighting, regarding its computation for the Turkey Demographic and Health Surveys, as well as its computation for the combined dataset. The calculation of sampling errors in complex DHS sample settings is explained after weighting; followed by an explanation on how they are computed in the combined data set. The last sub section of the data and methods section is on the selected variables. Six different indicators are described in this section. Section 4 is the Results section, presenting the results for means, proportions and rates (all six selected variables). Results include combined data estimates of these variables. The discussion is found in Section 5, where the findings are interpreted. The thesis ends with a conclusion section (Section 6).

2. LITERATURE REVIEW

Given this dissertation has a major methodological dimension, there is a wide spectrum of literature to be considered. This section focuses first on the development of demographic sample surveys, and then gives examples of rationales behind combining data sets. It will then move on to applications on combined data sets, including pseudo-panel methods and regressions with time dummies.

2.1. Development of Sampling Theory

The idea behind sample surveys is to collect data on a subset of a population to make inferences about the whole population. The process results in what is called sampling errors; any inferences reached by a sample will be just estimating the real population values with a degree of precision that depends on a number of factors like sample size. Its advantage comes from its low-cost compared to surveys, and the potential to collect more detailed information.

Brunt (2001) traces sample surveys back to Arthur Young, who conducted the very first survey of people, farms and villages, back in the 18th century. Yet it was not until Anders Nicolai Kiær became the Director of the Bureau of Statistics in Norway did sampling theory began to be systematically developed (Seng, 1951). He introduced a term called “dénombrements representatives” (representative investigations) in 1895, meaning neither a haphazard enumeration, nor a full inquiry, but rather a particular representative method of selection – what is now known as purposive sampling (Seng, 1951). Arthur Lyon Bowley in England, had also been an advocate of the use of sampling since 1906, joining into discussions with the International Statistical Institute (ISI) where Kiær has also advised before (Seng, 1951; Aldrich, 2008).

Stating "…if no measurement of precision can be made, the results are valueless", Bowley proposed the use of the theory of probability in sampling (Aldrich, 2008; Seng, 1951). Unlike Kiær's representative method based on purposive sampling, Bowley introduced random sampling with equal probabilities (Bethelem, 2009).

Aymler Fisher's book dated 1923, The Design of Experiments was influential in terms of the actual use of randomization. A year later, International Institute of Statistics (ISI) appointed a committee to study "the application of the representative method in statistics", that came up with a report in 1926, which included purposive sampling and simple random sampling (Seng, 1951; Bethelem, 2009). Verjin Stuart, Director of Statistics Netherlands made some comments on the issue on the same year, advocating random sampling over purposive, the latter reflecting subjective decisions (Bethelem, 2009).

Jerzy Neyman criticized purposive sampling theoretically in his 1934 paper, proving it could produce unsatisfactory estimates of population characteristics, empirically. Perhaps more importantly, he carried theory that was shaped around point estimation to the next step by coming up with the concept of "interval estimation" (Seng, 1951). This paper was a pivot in from theory to application too, into real sampling plans for actual randomly selected large scale surveys (Fienberg and Tanur, 1966). In the meantime, Rothamsted Experimental Station in UK, the longest running agricultural research in the world, included a group of statisticians had also been working on various experiments, including Fisher, Mackenzie, Clapham, Irwin, Cochran, Wishart, Yates, Zacopany, the findings of which have been summarized by Yates in 1946 (Seng, 1951). Among these is the estimation of the precision of sample estimates solely on the basis of the information in the sample (Bethelem, 2009). Stratification was emphasized for the reduction of the magnitude of sampling errors, and multistage sampling was introduced for sampling in agricultural surveys (Bethelem, 2009).

The development and application of sampling theory shifted from Europe to the Unites States of America around the previous decade (Seng, 1951), not to mean it did not exist there earlier. It existed in a more applied sense: Carroll D. Wright, chief of the Massachusetts Bureau of Labor, has been using the "representative method" in early 20th century, which he is known to have designed haphazardly rather than randomly (Seng, 1951). Developments were also being made at the Department of Agriculture and U.S. Bureau of Census. William G. Cochran joined the Statistical Laboratory at Iowa State University in 1939, where Arnold King and Ray Jessen had already been working (Hansen, 1987). Cochran published papers on regression estimation in sample surveys and systematic sampling in 1942 and 1946 (Hansen, 1987).

In the meantime, studies on sampling were continuing in Europe and various other parts of the world. Statistics Netherlands carried out a first test of a real sample survey using random selection in 1941 (Bethelem, 2009). There were statisticians in Russia dedicated to the topic like A.V. Peshekhonov and Kovalevskiy even before 1930s (Berthelem, 2009). There was also an interest in India. Prasanta Chandra Mahalanobis, also the head of the Indian Statistical Institute in the 1930s, made notable contributions by introducing pilot surveys, and by his works on the concept of optimum survey design and interpenetrating network of samples (Rao, 1973; Hansen, 1987). In the international field, The United Nations Statistical Commission established the Sub-commission on Statistical Sampling in 1947 that included statisticians such as Darmois, Deming, Mahalabobis, Yates and Fisher; and published documents that would assist national statistical institutes (Bethelem, 2009). According to Bethelem (2009), the classical theory of sampling was more or less complete by 1952.

However, advances continued. Hansen (1987) lists three other important contributors to the field, who have not only developed theory, but have also trained many students in the field: Hartley, Leslie Kish and Tore Dalenius. Leslie Kish has worked in fields as the popularization of computing and analyzing sampling errors and design effects for complex samples, came up with a procedure to sample persons within a household (aka Kish table) and small domain estimation (Verma, 2001).

Ever since they became widespread, sample surveys have been used to collect information on a variety of different research areas; like agriculture, labor, trade, marketing, voting, demography, etc. Among sample surveys, probability samples gained popularity over time compared to purposive or quota sampling. The latter has been very popular among polling companies in the USA until the Truman vs. Dewey presidential election in the 1948, when they all failed to point out the correct winner. Survey Research Center at the University of Michigan on the other hand, surveyed a probability sample that did not reflect this common result, and the whole event was a turning point in survey methodology (Rosegrant, 2012), resulting in the promotion of statistical methods rather than ad-hoc procedures.

Sampling and survey issues are still being studied and the field is still expanding. With a growing number of applications if sample surveys, theory also grew, and a lot of detail has been introduced to the area of sampling.

2.2. Sample Surveys in Demography

Demographic survey waves have also benefited from and contributed to survey sampling. With the growing concern over population growth back in the middle of the 20th century, the measurement of the key component of the demography in terms of its effect on age structure-fertility, gained more importance than ever (Bjerve, 1973). Thus took place a variety of surveys in the developing world. Duncan (1973) classifies these first waves of surveys into four categories: multipurpose surveys, demographic surveys focusing on general enumeration, fertility surveys which measure fertility, and Knowledge, Attitude and Practice (KAP) surveys. KAP surveys have most commonly been considered as the precursor of the World Fertility Surveys (WFS). They were not coordinated or centralized, and there was no standard questionnaire or survey format (Halfon, 2007), and they were more concentrated on contraception than fertility. Examples of KAP surveys were carried out in the developing world, including countries like Turkey, Hong Kong, Singapore, Thailand, India, Sri Lanka and Iran.

These early demographic studies have served as a justification to the WFS series by showed the need for trustworthy, scientific and comparable studies in fertility (Halfon, 2007). In 1982, the International Statistical Institute launched the WFS as a large international project with the collaboration of the United Nations and the International Union for the Scientific Study of Population, involving the implementation of nationally representative and internationally comparable sample surveys on fertility (Bjerve, 1973; Verma *et al.*, 1980). The WFS covered 40 developing and 21 developed countries, using standardized questionnaires (Brainard and Zaharlick, 1992). Given it is not possible to define identical samples for each country, there were guidelines instead; advising a minimum and maximum number of respondents, self-weighted samples, listing, multi-stage area sampling, a fixed “take” (a fixed cluster size) and stratification (Verma, Scott and O'Muircheartaigh, 1980).

Contraceptive Prevalence Surveys (CPS) were an international effort, like the WFS series, but in terms of content, they were more similar to the KAP surveys (Lewis, 1983). This wave, as the name suggests, was more contraceptive use and family planning policy oriented than the WFS. Its initial wave was carried out by Westinghouse Health Systems, selected by the U.S. Agency for International Development (USAID). CPS surveys cost less and were finished faster, with smaller sample sizes compared to the WFS, lacking essential credibility by being viewed as sloppy, and inclined to USAIDS’s needs (Halfon, 2007). The importance of the CPS program is its influence on the next wave of demographic surveys, the Demographic and Health Surveys (DHS) Program (Halfon, 2007).

The DHS program was initialized in 1984, as a follow-on to the WFS and CPS programmes (Vaessen *et al.*, 2005), and it still continues today. The DHS program is funded by USAID. It was initially implemented by ICF/Macro, Macro International Inc., ORC Macro, Institute for Resource Development, Inc., Measure DHS and is currently implemented by ICF International. The DHS surveys, like the WFS, are surveys with certain standards: There is a sampling manual recommending on sample and survey design, there are core questionnaires that are updated occasionally, and ICF International personnel can be involved in implementation, data processing and tabulation. So far, 90 different countries have been involved in the DHS (DHS Program, 2014).

The surveys mentioned above are those that are relevant to the demographic survey tradition in Turkey. There are other multinational surveys such as the Gender and Generations Surveys (GGS), Multiple Indicator Cluster Surveys (MICS), European Social Surveys etc. as well as some country level surveys; such as the National Health Survey, the National Longitudinal Survey of Adolescent Health and the National Survey on Family Growth in the US.

2.3. Combining Surveys

There is a wide literature on combining surveys. There are different foci and aims, and some are more methodological than empirical. This section will review these by first presenting the conceptual issues, and then presenting some empirical examples; some that are closer to the issue on this dissertation, and some that are rather different.

2.3.1. Conceptual Issues

Combining surveys over time or domains for observing trends and comparisons, improving coverage, improving precision and measurement error are issues on which literature is still being built up. Leslie Kish was among the first to suggest the cumulation of surveys in his no classic book dated 1965, by mentioning time sampling (Kish, 1965). Small area estimation, some problems of which can be considered as a combining issue, started to develop in the 1970s and 1980s. Further examples of combining surveys became common from the 1990s and on. Kish (1999), on one of his final papers, was suggesting that combinations or cumulations of surveys over space or time must receive serious technical consideration in the future, since all such work was done by ad hoc procedures so far, also mentioning periodic surveys as one of the "new paradigms" for probability sampling (Kish, 2000). His proposed terminology was to use "combining" for integration over space, and "cumulating" for integration over time (Kish, 1999), however, since not all the literature adopted this use, we used the terms interchangeably in this dissertation. He mentioned the World Fertility Surveys in the 1970s and Demographic and Health surveys in the 1990s as multinational designs that could be used to combine data across countries, and discussed possible methods to do so. As for cumulations, he gave the examples of the Health Household Interview Surveys in the US, the Australian Population Monitors, and the new Labor Force Surveys of the UK. Since the former is not within the scope of this dissertation, only the latter will be focused on.

Kish (1998) argued that many people perceived a conceptual difference between temporal combination and spatial cumulations; that combining regional or provincial statistics into national aggregates was "natural", but cumulations to annual or decennial aggregates seemed unusual. He argued that this was probably more psychological and social than philosophical. He also underlined the importance of the time intervals involved when cumulating data, suggesting weekly aggregations are more readily accepted compared to decennial ones. The way he conceptualized time sampling or cumulating across time was regarding time intervals as strata: He suggested that separate periods of a repeated sample could be regarded as strata and domains of the entire survey interval (Kish, 1965).

The reason why Kish would advocate the cumulation of data over time was that this would strengthen the sample base domains, especially for small ones and for variables that are relatively less stable. He further argued that cumulations of cases could be more efficient than weighted combination of statistics for rare items and small samples (Kish, 1987).

A conceptually related issue to cumulating samples is a “rolling samples”, developed by Kish in a series of papers (Kish, 1979a; 1979b; 1981a; 1983; Kish and Verma, 1983; 1986; Kish, 1990; 1997; 1998; 1999). Kish (1999) argued that censuses provided spatial detail, but not temporal detail, and small-area estimation was only a partial solution. Moreover, although surveys were more frequent, they could not provide the spatial detail censuses did. Administrative records, on the other hand, could satisfy the need for both temporal and spatial detail, but their qualities were high in only a limited number of countries. His proposed solution, rolling samples, was the cumulation of data from periodic surveys that would roll over the whole population in a relatively long period, so that they could both provide frequent, timely data, and spatially detailed data when aggregated.

As mentioned by Kish (1999), the American Community Surveys (ACS) is an example of rolling samples. This survey replaced the census long-form questionnaire that was sent out to about one in six households, for the last time in 2000 (U.S. Census Bureau, 2009). It has a rolling sample design, providing annual data which used to be only available once per decade before. The annual data of the ACS is also cumulated for more detailed analysis of domains. However, the cumulation of annual samples in the ACS into multi-year averages (3 or 5 years) presents a tradeoff (Transportation Research

Board of the National Academies, 2007). It was argued that while sampling errors of estimates are reduced through increasing sample sizes, the interpretation of results would get difficult for areas that experience rapid change over the years.

Another issue to keep in mind for the analysis of multiyear ACS data was about setting a common, synthetic interview date and common geographic coding (Transportation Research Board of the National Academies, 2007). The interview month variable is recoded, so that the data from the entire period appears as if they came from the same one-year period. This is similar to the approach used in this dissertation, as will be explained in detail in the Methodology section. As for the consistency of geographic codes, the ACS updated the addresses in the sample with the newest geographic coding available, keeping the geographic definitions consistent throughout years.

There are also examples of combining surveys over time that were not initially designed for this purpose, unlike the ACS. The next section will provide examples for these.

2.3.2. Integrated Survey Series

There are several projects of combining sample surveys over time that were designed as independent national surveys in the US. The major ones are mentioned in this section.

The Integrated Public Use Microdata Series (IPUMS-USA) is a database of more than fifty high-precision samples of the US population, obtained from the long term questionnaires of the censuses and their successor, the American Community Survey (Minnesota Population Center, 2012). The project is run by the Minnesota Population Center at the University of Minnesota. The samples, being created at different times, are somehow different in terms of record layouts, coding schemes and documentation, and the IPUMS harmonized these as much as possible. All the IPUMS samples are cluster samples, and there is either explicit or implicit stratification. Yet IPUMS users are said to rarely estimate the true standard errors, because it is time and effort consuming. Being the result of the rolling samples of ACS, stratification is consistent, so strata codes are kept the same throughout the years (i.e. strata are not re-numbered over the years when combined). As for geographic codes, a variety of geographic variables were constructed for the IPUMS-USA, so that historical comparisons for geographical areas could be made.

The Integrated Fertility Survey Series (IFSS) is another dataset harmonization project, carried out by the Population Studies Center and the Inter-university Consortium for Political and Social Research at the University of Michigan (University of Michigan Population Studies Center and Inter-university Consortuim for Political and Social Research, 2012). The database includes data from 10 fertility related surveys, such as the National Fertility Survey and National Survey of Family Growth in the US. The project provides users with a single dataset that has harmonized variables across all surveys. Harmonization was required due to changes in research objectives, question wording, and even in terms of basic definitions of concepts like race or poverty. The survey universes also differed in time. For example, the 1950 Growth of American Families Survey was restricted to a subgroup of white women only, whereas the later surveys did not have any race criteria.

The IFSS surveys had stratified multi-stage cluster designs; and both strata and cluster variables are present in the data. The values for the stratification variable, unlike the ACS or IPUMS, are recoded to avoid overlap across surveys. This means, for a stacked dataset, if the strata values go from 1 to H for the first survey, it goes from H+1 to H+n for the next, and so on. This is done as a means of convenience for users who wish to use multiyear data. There is no such problem for cluster variables, as software handle clusters within strata.

The Integrated Health Interviews Series (IHIS) is another data harmonization and integration project, also carried out by the Minnesota Population Center (Minnesota Population Center and State Health Access Data Assistance Center, 2012). It includes data from the US National Health Interview Surveys (NHIS). The NHIS is a face-to-face HH interview that is carried out every year since 1957. It is the longest running health survey in the US. It has a complex, multistage design, with stratification, clustering, and oversampling of some subpopulations. The original data files were recoded so that they could be as comparable as possible, thus there are differences in variable names etc. compared to the original public use data. For example, questions regarding education in the NHIS changed over time, and the integrated data set includes a variable that standardized this information into a new variable. For questions whose response categories were wider in some surveys, response categories were reduced to the narrowest available. Documentation is available for all related variable recoding.

2.3.3. Demographic and Health Survey Examples

There are few examples of combining Demographic and Health Surveys Data in the literature. While some aimed to observe time trends by doing so, some aimed to improve precision. This section will first review these studies, given they are relevant to the dissertation topic; and then move on to other examples from other surveys to portray the variability in the issue.

Pullum (2008), in his methodological paper on data quality, used combined data over 81 surveys. Many countries were represented more than once in this large dataset. Because of the nature of the study, usually equal weight was given to each survey; employing the sample weights of each survey at the same time. For some of the analysis, weighting was ignored completely.

Rutstein (2008) pooled 52 DHS surveys across time and space to study the effects of birth intervals on early age mortality and childhood nutrition indicators. There are multiple datasets on certain countries in this large pooled dataset, such as Egypt, Armenia, Bangladesh, Colombia, Ethiopia, Malawi and Peru. Although there is a brief note in the inclusion of stratification and clustering, no specific details on sample eights are provided.

Retherford et al. (2010) pooled data across three consecutive Philippines DHS surveys carried out in 1993, 1998 and 2008. Their study suggested a multivariate analysis methodology to estimate the effects of socioeconomic variables on the parity-progression based total fertility rate; and the effects of these on the trends in this indicator. While they used cross-sectional data for the former, they used pooled data from the three surveys in the latter. They pooled birth history data as records of years; and for each parity, they created data sets based on both cohorts and periods from all three surveys. Therefore they could use both a period and a cohort approach to estimate the effects on trends in TFR.

Sneeringer (2009), in her comparative analysis of fertility trends in Sub-Saharan Africa, exploited DHS birth history data as if it were panel data. She transformed women’s birth history data into person-year level data, each case including women’s age and total number of births for the corresponding year. She presented cohort and period fertility measures by mother’s birth cohort in her findings. More than one survey data set has been utilized for countries when available.

Karatepe (2010), in addition to using economic data sources, pooled Turkey DHS data over three samples; 1993, 1998 and 2003 to perform analysis on infant mortality. The aim was to increase sample size and expand the covered period, so that a wider perception on the relationship between socioeconomic characteristics and infant mortality could be obtained. In this publication, no details were provided for the use of weights or variance estimation.

Kırdar et al. (2012), in their study on the effect of compulsory education on teenage marriage and childbearing, pooled two sets of Demographic and Health Survey data (TDHS-2003 and TDHS-2008) to increase the precision of their estimates. Sample weights in the data sets were used in the analysis, without additional modification. Complex sample properties were not mentioned in the text.

Koç and Eryurt (2013) pooled data from four different Turkey Demographic and Health Surveys (TDHS-1993, TDHS-1998, TDHS-2003 and TDHS-2008). They used the pooled data set to analyze the relationship between consanguineous marriages and infant deaths. They used an adjustment factor for sample weights as suggested by Mariott et. al. (2007) that gives an equal weight to each data set. Other complex sample properties (i.e. stratification and clustering) were not utilized in the study.

2.3.4. Other Examples of Combining Surveys

There is a variety of reasons for combining data across different surveys, with the common aim of manipulating data into further use than was originally planned.

Schenker et al. (2002) combined data with the aim of **improving coverage**. They used data from the National Health Interview Survey (NHIS) and the National Nursing Home Survey in the US. The main interest was the prevalence of chronic disease, and the NHIS was a household survey. The inclusion of nursing homes in the estimate was crucial especially for the prevalence of chronic disease for the elderly, so the data was combined to represent the populations of both households and nursing homes at the same time. Their estimation method was the treatment of the two surveys as if they were separate design strata, and carry on with complex sample data analysis. Their statistic of interest was a ratio, and they used Taylor series linearization for variance estimation.

Schenker and Raghunathan (2007) give an example of using a household survey and US census for **providing comparability**. The Office of Management and Budget in the US changed the reporting of race in 1997; while only one race could be reported prior to this date; multiple races were available as responses after this date. For computing vital rates, The National Center for Health Statistics would obtain numerators from state level systems; and these systems would use the older standard for race classification. The denominators, on the other hand, would be obtained from the US Census, which would use the new classification. This would make it impossible to calculate rates for race groups. The National Health Interview Survey included both a question that would allow multiple race reporting, and a question that makes the respondent pick only one answer. The approach taken by the National Center for Health Statistics was to produce estimates of the denominators that would be observed if the old race classification was in effect; and use bridging models from the National Health Interview Survey on the data to distribute multiple-race reporters to what would have been in the case of single-race reporting.

Dong (2012) in his doctoral dissertation combined three different complex sample surveys with the challenge of **comparability**. His goal was to estimate 2006 insurance coverage from the Behaviour Risk Factor Surveillance Survey (a national telephone survey), National Health Interview Survey (a national HH interview survey) and Medical Expenditure Panel Survey (a survey including two separate surveys: one with a HH survey supplemented by data from their medical providers and another with a survey of employers that provide data on employer-based health insurance). He used a Bayesian bootstrap, which involved the generations of synthetic populations for variables that are unobserved in one survey, and observed in the others. These generated samples “uncomplexed” the two samples, meaning that the data were transformed into a self-weighted data set. Therefore they could be treated as thought they were from a simple random sample (Zhou *et al.*, 2012). The reason for the uncomplexing was the different sampling and non-sampling properties of the surveys.

A paper by Raghunathan (2006) includes an exercise on combining information from an interview survey and a health examination survey to improve the analysis of self-reported health questions, so that more **accuracy** could be gained. The National Health and Nutritional Examination Survey (NHANES) was used to assess the extent of error in the self-reported disease status, and a Bayesian framework was used to develop corrections for self reporting in the National Health and Interview Survey (NHIS). Measurement error models were fitted from the examination survey to the interview survey to adjust for potential inaccuracies.

Rendall et al. (2006) carried out a study to **improve the quality** of male birth rates, arguing that survey data on fertility on men are generally of low quality, and birth registration data for men is available in little detail. They estimated male fertility rates from the 2002 National Survey of Family Growth and evaluated them against those obtained from registration data. They proposed a Bayesian framework for allowing the incorporation of population data and women’s data into the results.

Hassett et al. (1995) combined data to produce **timely estimates**. Their problem was that despite the monthly data collection of the Energy Information Administration surveys in the US, estimates of oil prices would not be available before 2 months due to processing time and the nature of the surveys. Therefore they utilized daily spot prices together with survey data to forecast oil prices. Transfer functions were used to relate the two different sources.

Michaud et al. (1995) worked on combining administrative data and survey data to **reduce respondent burden**. Their application was on the Survey of Labour and Income Dynamics in the US, which is a panel survey. The idea was to do a micro-match and link data to income tax files from administrative records, with respondent consent. Having received only a few consents, it was decided that respondents would be offered either a questionnaire or access to their tax files.

Nusser et al. (2004), in their agricultural study, combined data from state and national monitoring surveys to **assess policy impact**. Their sources of data were the U.S. Department of Agriculture National Research Inventory (NRI) and Iowa Annual Pheasant Roadside Survey. They combined the information on a defined land level, obtaining land/cover use from the former, and count data from the latter, and used the information for assessing the impact of the Conservation Reserve Program on pheasant populations.

There are many studies that aim to improve small **area estimates**. Raghunathan et al. (2007) used the Behavioral Risk Factor Surveillance System (BRFSS); a large, state based telephone survey, and the National Health Interview Survey (NHIS); an area probability sample face-to-face survey. The sampling frames, designs and response rates varied among these two surveys. A hierarchical Bayesian approach was used to combine the date and address nonresponse and coverage errors. The authors obtained county level small-area estimates from the combined data for smoking, mammography and pap-smear rates through this methodology.

Qian and Kaplan (2001) and Qian et al. (2002) aimed to provide more accurate estimations, especially for **small groups**, using data from two separate samples. They analyzed design effects and performed statistical power analysis for a combining surveys application, in a setting where a national sample and state samples co-exist for the same survey: National Assessment of Educational Assessment (NAEP) in the US. They created a set of shrinkage weights for this purpose, which allowed mean statistics to have minimum variance estimates. They found greater design effects due to decreasing efficiency, yet ended up with smaller variances due to larger sample sizes (Qian and Kaplan, 2001), as there are larger cluster effects for the combined sample (Qian et al., 2002).

There are many examples in the literature for **small area estimations**. Ericksen (1973), in his demographic study, suggested a new method of postcensal estimation, after the vital rates technique, Census Method II, the composite technique and the ratio-correlation technique were developed. The method combined symptomatic data with sample data from the Current Population Survey by means of a regression model. The aim of this was to provide population estimates for local areas. Symptomatic data was available in the form of registration data on birth, deaths and school enrollment, which are roughly parallel to the changes in the size of total populations. The method linked information in PSU level.

Zieschang (1990) examined some regression-based methods of survey weighting that could be used to provide a “best” estimate from two combined surveys. He discussed the Principal Person weighting method (post-stratification adjustment that updates data according to age, race and sex categories) and the generalized least squares (GLS) method. He based his empirical study on the Consumer Expenditures Survey in the US, which included two different modes of data collection with the same target population and sample frame; one being a diary, and the other being an interview component. He found that GLS is useful because it allows the linking of multiple surveys easily. He further suggested that it provided population estimates with better precision compared to other methods.

Renssen and Nieuwenbroek (1997) generalized Zieschang's (1990) method. Their problem was to provide the same population totals from common variables of two surveys, where common variables were defined as variables that were present in both surveys, but whose population totals were unknown. Their method is based on post-stratification. The authors suggested that reaching the known population totals through post-stratification can also be achieved by using the post-stratification variables (age, sex, race, etc.) as regressors in the general regression estimator. Therefore they proposed the use of common variables as additional regressors in the pooled data set. The implicitly defined weights if the adjusted regression estimator would reproduce the estimates of the population totals for the common variables. They proposed a constrained minimum distance method to align the estimates of comparable totals among the two surveys, which reduces to Zieschang's (1990) under certain conditions.

Wu (2004) worked further on Zieschang's (1990) and Renssen and Nieuwenbroek's (1997) methods, arguing that the weights associated with their generalized regression estimator with auxiliary variables can turn out to be negative. The author suggested that his approach had straightforward interpretation for maximum likelihood, and his weights were always positive. His approach was the use of pseudo empirical likelihoods.

Chand and Alexander (2000), in their study on the American Community Survey (ACS) focused on multi-year averages from the rolling samples of the survey. While the ACS would produce reliable annual estimates for their indicators of interest with a population of 250,000 or more, smaller areas would require the cumulation of multi-year data. They wondered how many years of data should be employed for this, and how this would differ than a single-year snapshot. Their criterion was to obtain sampling errors to the multi-year averages that were close to those of the annual ACS estimates, predicting a five-year average. In their empirical application, they compared three year averages to five year averages, and found that longer averages means larger reductions in the mean square error and they performed better at smoothing noise and spikes.

Merkouris (2004) worked on a generalized regression procedure to combine data from multiple surveys, or multiple samples from the same survey. Common survey characteristics were employed as additional auxiliary variables to align estimated totals. He proposed a modification of Zieschang’s (1990) method to account for the differential in effective sample size between two surveys, claiming to improve the efficiency of the derived composite regression estimators. In a later study, Merkouris (2006), suggested a method of small-domain estimation that was design-based, as opposed to the existing model-based methods. He combined data from multiple surveys through regression, with the aim of borrowing strength from other surveys of the same population that have collected comparable information on variables of interest. He suggested that combining improves the precision of domain estimates, based on his empirical application to Canadian Labor Force Survey rotation samples which had stratified multistage are designs and were independent. Merkouris (2010) later worked on an adaptation of his work in 2004, with some options for incorporating auxiliary survey information in the estimation procedure. He proposed a regression method that was an extended calibration process, in which domain estimates from multiple surveys were calibrated into each other.

Ybarra and Lohr (2008) suggested a method of small estimation in which auxiliary data available could also be subject to measurement error, i.e. surveys. They based their work on the Fay and Herriot (1979) model, which would employ auxiliary data from registration type data with no sampling error. Ybarra and Lohr (2008) worked on two sample surveys. For estimating the health characteristics measured in the NHANES, they used data from the NHIS as auxiliary variables, whose sample size was almost triple that of the former. Instead of working with population quantities, they worked with population estimators and used measurement error models.

2.3.5. Weighting Issues in Combining Surveys

As would arise with spatial combinations, issues of weighting arise with temporal cumulations. This section reviews the literature on different weighting schemes used when combining sample surveys. It will begin with Leslie Kish’s approach to weighting for cumulations, then go on to weighting in the American Community Survey, an example of rolling samples. Then, examples of data integration projects in the US will be provided (IPUMS, IFSS, IHIS, SESTAT), other examples from US and Canada will be reviewed, and the section will conclude with two DHS studies. The American tradition is using weights that add up to the whole population, and are adjusted with post stratification. Demographic and Health Surveys on the other hand, employ weights that sum up to the total unweighted sample size and no post stratification is used. Thus while the main idea is usually maintaining the same population size on the former, it is maintaining the total unweighted number of cases in the other.

Kish (1992) listed seven main reasons for using weighting; ranging from disproportionate sampling fractions to frame problems. Among them was combining samples, which he said was becoming more popular, important and feasible, because of the increase in samples that could be combined. All combinations, in his view, concerned weighting in some form. He suggested

(Kish, 1999) that while cumulations of data with the same sampling fractions over a year could be useful for averages, a 10-year average would be affected by secular trends, thus secular trends in population size must be considered.

In his discussion on weighting issues for cumulated samples, Kish (1999) mentioned four alternatives on an example of combining 10 annual samples:

$\bar{y}_c = \bar{y}_0$, with $W_0 = 1$, and all other $W_i = 0$, assigning the full weight to the base year,

$\bar{y}_f = \bar{y}_9$, with $W_9 = 1$, and all other $W_i = 0$, assigning the full weight to the final year, providing the most timely estimate,

$\bar{y}_e = \sum W_i \bar{y}_i$ with $W_i = 0.1$, equally weighting all samples, which may be good for changeless stability,

$\bar{y}_w = \sum W_i \bar{y}_i$ with $W_0 \leq W_1 \leq W_2 \ldots \leq W_8 \leq W_9$, with monotonically nondecreasing W_i, where the curve of the increase may be determined by some model or empirical data, or both. He argued that a monotonic increase like $W_{i+1} = W_i + k$ (with $k > 0$), or even better $W_{i+1} = cW_i$ (with $c > 1$), seemed better than $W_i = 0.1$.

The American Community Survey (ACS) includes both person weights and household level weights for the following rationales: (1) differences in sampling rates across different areas (i.e. different selection probabilities), (2) differences between the full and interviewed samples (i.e. non-response error) and (3) differences between the sample and the independent estimates of basic demographic characteristics (i.e. post-stratification) (U.S. Census Bureau, 2009).

When the annual samples of the ACS are cumulated into multiyear datasets, a procedure similar to the single-year weighting methodology is used (U.S. Census Bureau, 2009). For multi-year averages, data are pooled together into one file. Weights are adjusted by the inverse of the number of years in the combination period, so that they contribute proportional to their share. For example, the base weights are divided by three when three year estimates are required. An additional step is post-stratification on this final weight, so that it matches the average population counts for these years, obtained from time series. A final step is model-assisted estimation, which is required for lowering the variances of estimates for tract-level estimates. Model assisted estimation involves six steps: (1) the creation of frame counts for required geographic areas, (2) linking of administrative records to ACS frame, (3) forming unweighted geographic totals of the linked administrative record characteristics, (4) application of *weight after the mode BIAS factor* (i.e. bias arising from response levels to different survey modes) weights to the linked administrative records (household level) that fall in the ACS sample, (5) fitting of a generalized regression estimation model to calibrate ACS weights (so that weighted totals of linked ACS records to match the unweighted total from step 2 and the weighted number of housing units from the ACS matches the frame totals in step 1), (6) Proceeding with the remaining steps of ACS weighting (the details of single year weighting can be found in US Census Bureau (2010)) starting with *housing unit post-stratification* adjustments.

Most samples in the **Integrated Public Use Microdata Series** (IPUMS-USA) database are unweighted “flat” samples, meaning each individual or HH represents a fixed number of units in the US population (Minnesota Population Center, 2012). There are three different weights in the harmonized dataset that users can use: a household level weight, a person-level weight and weights specific to the samples that include sample-line characteristics (a criterion for selecting a subset of the population for the long form census in some of the datasets).

Replicate weights for IPUMS are provided for complex sample analysis in the database, and a project is currently underway (IPUMS Redesign Renewal) for developing more user-friendly ways of variance estimation under complex sample settings. Multiyear data sets (3 or 5 years) for recent years are currently available (they are also the datasets prepared for ACS) for researchers who wish to study small populations, in which weighting variables yield estimates for the entire 3-year or 5-year period, instead of single year populations.

In the **Integrated Fertility Survey Systems** (IFSS) project, over and under-sampling and non-response differed among the harmonized surveys. There were also samples that did not have sample weights available in data sets, and the IFFS team developed these for users. Post-stratification adjusted weights are also available in the datasets. Although the use of these resulting base weights is strongly recommended for analysis from a single survey, users are warned against regarding them as a satisfactory weight for analyses that combined or pools data across surveys.

The IFSS team regard the use of multiple years of data as a process of sampling over time periods like period prevalence estimates, because rates computed over pooled data gives proportions over a specific time period. Since the single survey weights add up to the whole population in the IFSS data, when multiple years are pooled, the population is represented by as many number of times as the selected number of surveys. If averages or proportions are computed, no adjustment suggestion is made for weights; however, it is required for totals. There are two approaches mentioned. One could either divide the result by the number of surveys used, or adjust the weights in advance so that the resulting total would be an average of the surveys used. The assumption here would be that all surveys would be equally spaced in time.

For the **Integrated Health Interviews Series**; sampling weights need to be used due to the presence of oversampling in NHIS data (Minnesota Population Center and State Health Access Data Assistance Center, 2012). There are four steps for weighting in this survey: (1) the inverse of selection probabilities, (2) household level non-response, (3) an adjustment for the correction of potential bias due to sample under coverage, (4) a post stratification adjustment for age, sex and race according to US Census Bureau population counts. Since the sum of weights yield the annual population in these surveys, weights need to be adjusted when data is pooled over time. The project team suggests the simplest solution to be the adjustment of the weight by dividing it by the number of years involved, and they mention users may also use more sophisticated methods if required.

A more recent documentation on NHIS (Center for Diesase Control, 2013) which is prepared from the survey description documents for 2006-2012 gives three classifications into which pooled analysis can fall: 1) the years of NHIS to be analyzed fall into the same sample design period; 2) the years of analysis fall into different sample design periods; 3) the years of analysis fall into the same sample design periods, but there were changes to the public use design variables.

In the first case, it is underlined that it would not be correct to treat the different years as independent, because the samples were drawn from the same PSUs. The document states that a valid approach for this kind of data set would be to use it like a very large data set instead. For the second case – which resembles the case in this thesis – the correct method is to treat the samples as independent, thus re-numbering strata to make sure the software used can compute estimates according to these independent samples. For the third case, NHIS leaves it up to the user by presenting an option: The data user would have to treat samples as independent (and accept the fact that standard errors would be underestimated).

The **Scientist and Engineers Statistical Data System** (SESTAT) is another integrated data system including three independent surveys: (1) National Survey of College Graduates, (2) National Survey of Recent College Graduates, (3) Survey of Doctorate Recipients. Since the Survey Populations have some overlap in their target populations (the first can apply to the second or third, and so on), they linked the data in a way that each person had only a single chance of appearing in the combined dataset. They thus have two types of weights, one that is specific to each survey, and one that accounts for multiple selection probabilities (National Science Foundation, 2012). Jang and Sukasih (2004) worked on replicate variance estimation on this database. While conventionally replicate weights of each survey were separate, the authors compared this to using a common set of replicate weights for the combined dataset. They concluded by recommending this approach over the other, suggesting it provides better estimators, as well as potential savings in terms of time and money.

Chu and Goldman (1997) explained the weighting procedures in the U.S. Department of Agriculture's **Continuing Survey of Food Intakes by Individuals** from 1994-1996. Nationally representative samples of persons were selected for each of the three consecutive years. Although the paper primarily focuses on post-stratification adjustments through raking, it also discusses the time dimension. The authors suggest that if year to year changes in, say average food consumption, is negligible, then the annual samples can be pooled with the base weights attached, and non-response and post-stratification adjustments can be applied to the pooled sample weights. However, the combined estimate would be biased if the expected value is not constant over time. Their proposed solution was to re-calibrate non-response adjusted weights to the corresponding three year average Current Population Survey population counts. For estimating the mean of a variable z, the approach was $\bar{z} = \sum W_i z_i$, where $W_i = N_i / \sum N_i$, and N_i is the size of the population in year i. The estimates obtained with this recalibration would be applicable to statistics regardless of assumptions of changes over time. However, the authors warn that this new weight will cause larger variances for statistics whose expected value does not change over time.

Friedman and Jang (2002) worked on the quarterly **Adult Health Care Survey of Department of Defense Beneficiaries** (HCSDB) in the US and looked into the issue of cumulating four of them into an annual dataset. The authors discussed several techniques on weighting for cumulated samples. One of the methods was the equal weights approach. The authors argued that this approach had the assumption that variation from one quarter to the next is only due to sampling variation, meaning the population estimate is assumed to remain stable over time. Therefore combined estimates could be calculated from the four independent quarterly surveys by averaging their estimates, which would be more precise then the quarterly estimates, due to the averaging out of the variance across the quarters. The authors described this approach as simple to implement, and easy to interpret.

The second approach discussed by the authors was weighting proportional to the reference period. This approach took into account the possibility that more recent data may be more relevant to data users or policy makers, and gives larger weighs to more recent periods. When this happens, the difference between the estimates of two successive quarters is no longer a simple difference; it is rather a linear combination of estimates, making interpretation more difficult. For the HCSBD, all quarters actually included data over a year; respondents reported a one year's experience at each survey, resulting in a nine-month overlap between each quarter. The data was thus weighted in a way that the overall sum of weights equaled the total population count. Each quarterly survey had to meet this condition: $WCOM_{REF} = q_i \times w_{q_i}$, where $0 < q_i < 1$, and with the constraints of $q_1 + q_2 + q_3 + q_4 = 1$ and $q_1 < q_2 < q_3 < q_4$. They picked these weights as $q_1 = 1/10$, $q_2 = 2/10, q_3 = 3/10$ and $q_4 = 4/10$.

The final approach described by the authors was weighting proportional to each quarter’s domain size for domain-specific estimates, given the sizes of domains could vary across quarters. They concluded that the equal weights approach was appropriate for the data, given the results do not differ a lot from that of the other approaches. They only advised users to be cautious against the use of the reference period technique for some of the domains, because of observed time trends.

O'Muircheartaigh and Pedlow (2002) compared two different weighting strategies for the **National Longitudinal Study of Youth** (NLSY) 1997. Although this survey is designed as a panel study, the base year (1997) included two independent samples: one cross-sectional sample with an approximately equal-probability sample of HHs, and another supplemental sample that over-sampled Hispanic youths and non-Hispanic black youths. The approach used until this study was combining the samples through a sample-based weighting system, where the relative relationship of case weights remained the same (cases with large weights still had larger weights after weight adjustment). While weights from the first sample were multiplied by λ, weights from the second sample were multiplied by $1-\lambda$. For an overall estimate of the statistic θ, the estimator was as follows:

$$\hat{\theta} = \lambda\hat{\theta}_{CX} + (1-\lambda)\hat{\theta}_{SU}$$

and the optimum λ was defined to be proportional to the relative effective sample size in the first sample:

$$\lambda = \frac{n_{CX}/d_{CX}}{n_{CX}/d_{CX} + n_{SU}/d_{SU}}$$

and

$$(1-\lambda) = \frac{n_{SU}/d_{SU}}{n_{CX}/d_{CX} + n_{SU}/d_{SU}}$$

where n_{CX} and n_{SU} are the nominal sample sizes and d_{CX} and d_{SU} are general factors (one plus the squared coefficient of variation of the weights within each sample, aka the design effect due to weighting: 1+L) used as the design effects. The design effects were calculated as:

$$\hat{d}_{CX} = 1 + [CV(W_i \epsilon\, CX)]^2 \text{ and } \widehat{d}_{su} = 1 + [CV(W_i \epsilon\, SU)]^2$$

The λ values were calculated separately different race/ethnicity and sex combinations, and the final weights were obtained by multiplying the original weights by λ or $1(-\lambda)$, depending on the sample and race/ethnicity-sex combination.

The new approach suggested by the authors, on the other hand, uses a Horvitz-Thompson approach, and re-weights each case, including the selection probabilities from both samples, i.e. obtaining the probability of being selected from either of the samples. These were calculated as follows:

$$W_i^* = \frac{1}{P_i^{CX} + P_i^{SU}}$$

where P_i^{CX} was the selection probability for case i in the CX sample, and P_i^{SU} was the selection probability for case i in the SU sample.

It should be noted that neither of these strategies involve non-response adjustments. The authors calculated the coefficient of variation of weights, the design effect due to weighting and the effective sample size, and concluded that their strategy was superior. They further suggested that it was also conceptually superior, because it ensures that cases with the same overall selection probability will receive the same weights. Starting with the 4th round of the survey, the weights started to be calculated with the second approach described in their paper.

Thomas et al. (2007) reviewed the weighting procedures used in the **Canadian Community Health Survey**. The survey differed from Kish's idea of the rolling sample, because it did not roll over the whole population, rather a large sample. Because of this nature of the survey, the authors regarded the estimated obtained from this survey as period estimates that would represent the evolving population over a time period. They stated that they gave and would continue to give equal weights to successive surveys when combining them across time.

Thomas and Wannell (2009) also discussed the methods of combining cycles of the Canadian Community Health Survey (CCHS), for researchers who may be interested in studying rare populations. The Canadian Community survey consists of two separate cross-sectional surveys, one is carried out in a regional level and provides data on general health characteristics, and the other one is carried out in a provincial level, providing more detailed health information. The authors began by a warning regarding the combining of these two different surveys, due to comparability issues. They suggested two different approaches to combining data from cross-sections; (1) the separate approach, (2) the pooled approach. The former involves the estimation from separate data sets, and then averaging them out. The latter is the pooling of data into one single dataset; regarded as an attractive option by the authors due to the power of the larger sample size. The key to implementing this is to keep the sample weights in the dataset, as well as the bootstrap weight files, which are necessary for the calculation of complex sample variances. Given the sample sizes add up to the population size in this survey, the pooled dataset would not perform for estimating totals, unless the weights are rescaled by a factor of α_i. The authors suggest this factor to be a constant one, such as $\alpha_i = 0.5$ if there are two surveys, or $\alpha_i = 0.33$ if there are three, because the assumption that each CCHS cycle can be used to estimate the same population parameter is questionable.

Thomas and Warnell (2009) further discuss that adjusting the weights is not always necessary when pooling data, based on the argument that resulting combined weights represent an average population that does not exist. Estimates of ratios, proportions and means do not change depending on the weights being simply pooled, or multiplied by a constant factor. The authors point regression analysis to be one of the main applications of pooled data, because more detailed models can be studied with a larger sample size, and time effects can be controlled for, if significant. The main conclusion of the study is that estimates resulting from combined CCHS samples are of higher quality than those from one cycle alone, yet users should consider the implications of the artificial population used in their study.

Retherford et al. (2010), in their study on **Philippines DHS** data of 1993, 1998 and 2003, renormalized the weights for analysis, so that the sum of weights would not exceed the total unweighted number of cases, sticking with the standard DHS approach. In notation;

If N is the number of unweighted person-year observations in the data set; w_i is the weight attached to the ith person and $W$is the sum of $w_i$over the person-year observations in the dataset; then the renormalized weight was calculated as:

$w_i^* = w_i(N/W)$, so that the overall sum would be equal to N. As for other compex sample properties; Jackknife repeated replications method was used estimate the standard errors of the complex measures calculated, like the TFR. The cross-sectional analysis included as many replicates as the PSUs on one survey, and the trend analysis included as many replicates as the total number of PSUs on all three surveys. There is no information on whether stratification was taken into account or not.

Sneeringer (2009), on her **DHS** based study on **Sub-Saharan African** fertility, used a weighting scheme that is in a way, similar to the approach used by Retherford et al. (2010); i.e. re-scaling the person-year level weights so that they would sum up to the overall unweighted number of cases. However, doing so, the author also recognized the relative share of the sample sizes of consecutive, same country DHS surveys. Scaling factors were calculated to make sure cases coming from larger samples were not over-represented in the estimates. This was done as follows:

If i denotes the individual, D denotes the index to the DHS in country C, and z is the total number of surveys in country C, the sample weight for the individual is denoted by $SW_{i,D,C}$. The sample size for DHS survey $D$is denoted by $SS_{D,C}$. Within the same country, the author argued, the relative DHS survey sizes should be taken into account. The basic idea was, if one sample is smaller than the other, and two cases had the same weights, the one from the smaller sample actually would actually represent more cases in the population. So a scaling factor, $F1_{D,C}$ was calculated as:

$$F1_{D,C} = \frac{\sum_{D=1}^{z} SS_{D,C}}{SS_{D,C}}$$

When this scaling factor is multiplied by $SW_{i,D,C}$, then the representation of the cases in the panel will be better in terms of selection probabilities. However, when this adjustment is made, the sum of weights no longer add up to the total unweighted sample size, so a final step was followed as:

$$F2_{A,C,Y} = \frac{n_{A,C,Y}}{\sum_{i=1}^{n} SW_{i,D,C} \times F1_{D,C}}$$

where A is age group and Y is year. Multiplying the original $SW_{i,D,C}$ with the two factors, one gets $W1_{i,D,A,C,Y} = SW_{i,D,C} \times F1_{D,C} \times F2_{A,C,Y}$, the final weight in the analysis (an empirical example can be found in Sneeringer, 2009). Further effort was taken on to combine data from multiple countries, the details of which are not presented here.

The final document to mention on weighting is the most recent DHS document on listing and sampling (ICF International, 2012). There is a separate section in this manual called “de-normalization of standard weights for pooled data” focusing on combining DHS data.

Normalization in DHS surveys is the procedure of making the weighted number of cases to the unweighted number of cases for both the household data set, and the women’s data set. As will be further explained in Section 3, the first step in the calculation of the sample weight in DHS surveys is to calculate the inverse of the selection probability; which is called the *raw weight*. This raw weight is then adjusted by non-response, and is *normalized* to match the number of completed interviews. This normalization is done as follows:

$$NW = \frac{1}{selection\,probability} \times \frac{\#\,completed\ interviews}{size\ of\ target\ population}$$

Where NW stands for normalized weight, the first term is called the raw weight, and the second term is called the *normalization factor*.

ICF International suggests the *denormalization* of the normalized weight by dividing it into the normalization factor. In order to do this, an estimate of the target population needs to be known, which can be obtained from external sources (census, registration, etc.).

The DHS manual suggests the use of either denormalized weights as explained above, or the re-scaling of the weights to the total number of cases in the pooled data set. It is also noted that this should be done separately for each set of weights available (i.e. household, women, men). The basic weighting approach taken up in this dissertation is in line with this DHS recommendation.

3. DATA AND METHODS

This section first provides the details of the data sets used for this dissertation. First of all, the sampling procedures, sample sizes and questionnaires used in the Turkey Demographic and Health Surveys (TDHS-1993, TDHS-1998, TDHS-2003 and TDHS-2008) will be explained. After this, the weighting adjustment that is used will be discussed, followed by individual sections on selected demographic indicators for estimation.

3.1. Data

Demographic surveys in Turkey provide direct demographic estimates which cannot be obtained from censuses or registration systems. Quinquennial cross-sectional surveys have been conducted by the Hacettepe University Institute of Population Studies since 1968, the last of which was TDHS- 2008. TDHS-1993, TDHS-1998, TDHS-2003 and TDHS-2008 have been implemented as a part of the Demographic and Health Surveys program. This program has implemented over 300 surveys in more than 90 different countries (DHS Program, 2014).

The Turkey Demographic and Health Surveys are very similar in terms of sampling design. They all have weighted, multistage, stratified clustered sample designs, i.e. complex sample designs. The sample sizes tend to increase in time due to renewed sampling needs, although not dramatically (Table 3.1). The similarities and differences between surveys are discussed in detail below.

3.1.1 TDHS-1993

Turkey Demographic and Health Survey, 1993 was implemented by the Hacettepe University Institute of Population Studies, in collaboration with the General Directorate of Mother and Child Health and Family Planning, Ministry of Health and Macro International (HUIPS, 1994). This survey consisted of two questionnaires: a household questionnaire and a women's questionnaire. While the former was filled for each selected household, the latter was filled for ever-married women aged 15-49.

The survey design of the TDHS-1993 allows for estimations for various domains as well as Turkey as a whole. These are urban and rural residences, and each of the five major demographic regions of the country: West, North, Central, South and East Anatolia. These regions were based on provinces, the number of which was 76 at the time. All settlements with a population above 10,000 were defined as urban. Moreover, district centers, regardless of their size were defined as urban, considering their administrative status. The frames were obtained from the 1990 Population Census, and the populations were projected for 1993 using the exponential growth formulas.

These five regions were divided into sub-regions for more detailed stratification. This was done according to the provincial infant mortality estimates obtained from the 1990 Population Census. As a result, 14 sub-regions were formed within the 5 regions. Each of these sub-regions were further stratified for urban/rural criteria, thus a total of 28 strata were created. Within urban strata, systematic selections of quarters were made with probability proportional to size. Quarters were divided up into segments of similar sizes, and these were sampled within quarters. Rural units were selected directly. A segmentation procedure was implemented during listing whenever necessary. All selected segments and rural units were listed to have up-to-date frames for the final stage of sample selection; and an average of 20 households were selected per segment for interviews. The number of segments included in the sample design was 500. However, 478 of these could be visited during fieldwork, due to problems of accessibility and security issues (HUIPS, 1994).

Table 3.1 Sample Implementation for TDHS-1993, TDHS-1998, TDHS-2003 and TDHS-2008

	Number of completed HH interviews	HH response rate	Number of completed individual (ever married women) interviews	Individual (ever married women) response rate
TDHS-1993	8619	96.8%	6519	95.0%
TDHS-1998	8059	93.8%	6512	92.7%*
TDHS-2003	10,836	92.9%	8075	95.6%
TDHS-2008	10,525	88.4%	7405	92.5%

Source: HUIPS, 1994; 1999; 2004; 2009

**Re-calculated by author for ever married women only for the sake of comparability.*

3.1.2 TDHS-1998

Turkey Demographic and Health Survey, 1998 (TDHS-1998) was carried out by the Hacettepe University Institute of Population Studies in collaboration with the General Directorate of Mother and Child Health and Family Planning (HUIPS, 1999). Both the United Nations Population Fund (UNFPA) and Macro International provided technical and financial support. Four questionnaires were used in this survey: A household questionnaire, an ever-married women's questionnaire, a never married women's questionnaire and a husbands' questionnaire. After the household questionnaire was filled, all women aged 15-49 were interviewed with different questionnaires based on their marital status (ever-married and never-married). A sub-sample of husbands of currently married women was also interviewed.

The sample design of TDHS-1998 is similar to that of TDHS-1993. The sample is designed to provide estimates for urban and rural residences and five main demographic regions. The sub-regions are the same as defined in TDHS-1993, resulting in the same total number of strata (28). Settlements were systematically selected from strata with probability proportional to size, and the measures of size were obtained from the 1997 Population Count. Among the selected settlements, those with municipalities had Structure Schedules available. Segments of approximately 100 households each were created from these schedules, and those were sampled at the next stage. For those settlements that did not have these schedules available, the approach was similar to that of TDHS-1993. The lists were created during listing. The complete settlement was listed if there were less than 250 households; and 200 were listed if it was more, and an estimate of the remaining number of households was provided. The cluster sizes were a fixed take of 25 in urban segments, and 15 in rural segments. The sample was designed to consist of 480 segments in total. Fieldwork yielded 476 segments, due to logistic reasons. The missing segments were either not visited because they could not be listed in the first place, or because they could not be visited after listing.

3.1.3 TDHS-2003

The TDHS-2003 carried out by the Hacettepe University Institute of Population Studies, in collaboration with the General Directorate of Mother and Child Health and Family Planning (HUIPS, 2005). Two questionnaires were used in the survey, like in TDHS-1993: A household questionnaire and an ever-married women’s questionnaire. The TDHS-2003 sampling design is different than TDHS-1993 and TDHS-1998, yet special effort was put in it to maintain comparability.

The major reason for the difference in sampling design was the adaptation of a new system of regional classification in 2002, in accordance with the accession process of Turkey to the European Union. According to the Nomenclature of Territorial Units for Statistics (NUTS), a new three level region classification was prepared for Turkey. The NUTS 3 level is the 81 provinces of Turkey, the NUTS 2 level is the aggregation of the provinces into 26 regions and the NUTS 1 level consists of 12 regions. These classifications are currently used for official statistics. The TDHS-2003 sample was required to provide estimates of certain statistics for the NUTS 1 regions, with special emphasis on Istanbul and Southeastern Anatolia regions. The adoption of the NUTS 1 regions and keeping the traditional five regions at the same time was a major challenge. Minor alterations were made to the geographic definitions of the five regions. Additional strata were created around regional boundaries, since the 12 regions cannot be aggregated to make up the five regions. Istanbul was divided into slum and non-slum regions with the co-operation of UN-Habitat, for an add-on study. Furthermore, estimates were required for areas which were affected by the 1999 earthquake in Turkey. All these considerations increased the number of strata to 40, 12 more than the previous survey.

For the selections of settlements within strata, measures of sizes were obtained from the 2000 Census of the Population. As was done in the TDHS-1998, Structure Schedules were utilized for settlements with municipalities, some of which were dated 2000, and some of which were updated in 2002. For Istanbul, the selection was on the basis of quarters, classified by expert opinion on slum/non-slum characteristics. For settlements without structure schedules, addresses were obtained through listing. If a settlement had less than 250 households, all were listed; and if there were more, than the first 250 were listed and a note was made on the number of remaining non-listed households.

Like in the previous survey, a fixed take of 25 households were selected in urban areas, and 15 were selected in rural areas. The total number of segments in the TDHS-2003 sample design was 700; and 688 of these could be visited for fieldwork.

3.1.4 TDHS-2008

The TDHS-2008 was implemented by the Hacettepe University Institute of Population Studies. The beneficiary institutions were the T.R. Ministry of Health and T.R. Prime Ministry Undersecretary of State Planning Organization. Two questionnaires were used in this survey, as was the case in TDHS-2003. The sampling design of the TDHS-2008 was similar to TDHS-2003, without the extra arrangements on earthquake areas and slum/non-slum. Therefore the number of strata is slightly less (36 strata).

Settlements were systematically sampled within strata, with probabilities proportional to measure of sizes obtained from the 2007 Address Based Population Registration System. Address lists were provided by TURKSTAT for sampled settlements that were municipalities, and lists were obtained for the rest through listing; exactly as was done in TDHS-2003. TURKSTAT sampled segments of about 100 households from the former, to be updated through listing in 2008. Cluster sizes were the same as TDHS-2003; 25 households per urban segment, and 15 households per rural segment. Segment level non-response is the lowest among TDHSs for TDHS-2008; 633 out of 634 sampled segments were visited for fieldwork.

3.2. Weighting

A need for design weights arises whenever samples are deviated from *epsem*[1]. When a simple random sample is drawn, or units are selected proportionally from strata, it is not required. TDHSs are stratified cluster designs that are non-epsem, thus weighting is necessary for valid inference. Another type of weighting is used to correct or reduce bias from non-response. If non-response is especially high in a given stratum; it will be under-represented in the overall estimates; giving a higher weight to observations from this stratum will reduce bias. Lastly, weighting is required if post-stratification is done. Post-stratification ensures that the basic characteristics of the sample resemble the population it represents. These characteristics are usually basic demographic characteristics like age and sex (ICF International, 2012).

The first two types of weighting are applied in TDHS surveys. Below sections include the details.

3.2.1. Weighting in Turkey Demographic and Health Surveys

Weighting in Turkey DHS surveys consists of two components: Selection probabilities and non-response. The reason why selection probabilities are important is that not all units are selected with equal probability in Turkey DHS surveys; in other words, they are not self-weighting designs. Based on certain sampling needs, some domains are sampled with higher or lower probabilities than they would have been samples under simple random sampling. Such sampling needs arise from statistical requirements. For

---

[1] Equal probability sampling, where each unit is sampled with the same probability of selection.

example, for a given demographic indicator, there might be a minimum required sample size to obtain a certain level of precision in estimations. Oversampling from some domains would be plausible for this case. Similarly, a proportional allocation for a given domain may provide too large a sample size that provides little gain in precision, but causes a large loss in terms of cost and time. This calls for undersampling.

In TDHSs, oversampling is usually observed for special domains (e.g. Earthquake strata in TDHS-2003) and regions whose populations are relatively small (e.g. Strata in the North region). Undersampling is mostly observed of domains with large sample sizes (e.g. Strata in the West region). The weights for sampling probabilities are called design weights, and are applied to make the sample resemble the target population (ICF International, 2012). The sum of the inverse of the sampling probabilities gives the total target population. In DHS, these are calibrated to give the total sample size.

The other component of DHS sampling weights, as mentioned, is non-response. Different units are known to have different response behavior. For example, urban households are less likely to respond than rural households (ICF International, 2012). The non-response weights correct for these, in an attempt to reduce bias.

Since the design weights are the inverse of the selection probabilities, the following paragraphs will formulize these first. All TDHSs, as mentioned above, are multi-stage: A selection of PSUs (settlements), followed by SSUs (segments) and then households (the fixed takes of 15 or 25).

The first stage selection probability is the selection of a PSU i (a settlement, denoted by M_{hi}) from stratum h, where n_h is the number of selections to be made from stratum h:

$$P_{1hij} = \frac{n_h \times M_{hi}}{M_h}$$

The second stage is the selection of a segment j (denoted by M_{hi}) of about size 100 from PSU i.

$$P_{2hij} = \frac{M_{hij}}{M_{hi}}$$

The final stage is the selection of a fixed cluster size m_{hij} from the selected segments, where $M_{hij}^{'}$ is the segment size after listing:

$$P_{3hij} = \frac{m_{hij}}{M_{hij}^{'}}$$

The overall selection probability is:

$$P_{hij} = \frac{n_h \times M_{hij} \times m_{hi}}{M_h \times M_{hij}^{'}}$$

The selection probability used in the sampling weight calculations of TDHSs assume that the fraction $\frac{M_{hij}}{M_{hij}^{'}}$ cancels out, leaving the overall selection probability to be:

$$P_{hij}^{TDHS} = \frac{n_h \times m_{hij}}{M_h}$$

The subscripts i and j are no longer necessary, because m_{hij} is determined by h. Thus the overall selection probability can be re-written as:

$$P_h^{TDHS} = \frac{n_h \times m_h}{M_h}$$

These fractions are presented at the appendices of each TDHS report. TDHS-1993 is somehow exceptional in terms of sampling weights; because it is the only TDHS that has does not have a separate weight for each stratum, i.e. the number of unique weights is less than the total number of strata. There are 11 different sampling weights in TDHS-1993 despite the presence of 28 strata. The reason is that selections in rural and urban areas were made proportionally, making these units self-weighting. Different selection probabilities were used for sub-regions in the East and Ankara in Central Anatolia, resulting in this number.

The design weights are thus calculated as:

$$W_h^{design} = \frac{1}{P_h^{TDHS}}$$

The design weights for household and individuals are the same, because no further selection is made within the household; all eligible women are interviewed with a probability of 1. Household and individual level weights differ only due to non-response adjustment. The TDHS non-response adjustments are calculated in a similar manner to the sample implementation calculations presented at the appendices.

The household response rates are calculated as:

$$HRR = \frac{C}{C + HP + P + R + DNF + PC} = R_{hh}$$

Where C stands for the number of completed household interviews, HP stands for no eligible household members present, P stands for postponed, R stands for refusal, DNF stands for dwelling not found and PC stands for partially completed.

The individual response rates are calculated as:

$$EWRR = \frac{EWC}{EWC + EWNH + EWP + EWR + EWPC + EWI + EWO} = R_{ph}$$

Where EWC stands for the number of completed individual interviews, $EWNH$ stands for not at home, EWP stands for postponed, EWR stands for refusal, $EWPC$ stands for partially completed, EWI stands for incapacitated and EWO stands for other.

The final sampling weights for households and individuals are as follows:

$$W_h^{HH} = \frac{1}{P_h^{TDHS}} \times \frac{1}{R_{hh}}$$

$$W_h^{I} = \frac{1}{P_h^{TDHS}} \times \frac{1}{R_{hh}} \times \frac{1}{R_{ph}}$$

It should be noted that this weighting scheme follows that of TDHSs, and Macro International (1996); and not that ICF International (2012); where the response rates are weighted by design weight.

3.2.2. Weighting in the combined data set

When using pooled data, using the original sample weights for each survey would result in a weighted statistic by the sample sizes of the surveys. Thus the effect of this on combined data estimates would depend on how different the independent sample sizes are. For instance, if the sample size of a survey is double the sample size of another survey, the former would have twice as large a weight in the final estimate. This, unless required for some special purpose, would result in estimates biased towards the survey with the larger sample size. The alternative used in this thesis is one similar to what was suggested by Kish (1999), so that the weights would reflect the secular trends in population size. In other words, instead of changing randomly, sample sizes would follow the increasing trend Turkey's population showed between 1993 and 2008.

In order to create such weights, the sizes of the target populations at each average survey date were calculated. These populations were obtained using exponential yearly growth rates and projected midyear population figures. Then all four sets of weights were adjusted in a way that they would add up to those population sizes. As a final step, the weights were calibrated to add up to the total sample size of all four surveys. Below is the formulation of this adjustment:

${}_{x}N_{5}^{it*}$ is the midyear population of the age group x to $x+5$, at year i when the survey was conducted. Here, $t*$ is used to denote the date of the population estimate – the midyear. ${}_{x}N_{5}^{it}$ is the population of the age group x to $x+5$ at the average date when the survey at year i was conducted, thus t is used to denote the survey date. The exponential growth rate is calculated between the midyear populations at year $t*$ and $t*+1$, denoted by $(r^{t*,t*+1})$ in the equation.

$${}_{x}N_{5}^{it} = {}_{x}N_{5}^{it*} \times e^{r^{t*,t*+1} \times (t-t*)}$$

Table 3.2 shows the midyear populations from 1993 to 2009 by age groups (${}_{x}N_{5}^{it*}$ values). This table was used to calculate yearly growth rates so that the populations could be adjusted for the median date of the survey. Table 3.3 shows the calculated growth rates $(r^{t*,t*+1})$; and Table 3.4 shows the populations at median survey dates (${}_{x}N_{5}^{it}$).

Table 3.2 The midyear populations of women aged 15-19 for selected years between 1993 and 2008 (x1000)

	1993.5	1994.5	1998.5	1999.5	2003.5	2004.5	2005.5	2008.5	2009.5
15-19	2 858	2 938	3 237	3 202	3 066	3 051	3 035	2 985	3 021
20-24	2 657	2 699	2 845	2 922	3 228	3 195	3 163	3 059	3 053
25-29	2 354	2 418	2 641	2 685	2 835	2 915	3 004	3 218	3 199
30-34	2 168	2 222	2 336	2 405	2 630	2 688	2 725	2 824	2 924
35-39	1 866	1 934	2 146	2 187	2 323	2 373	2 437	2 616	2 666
40-44	1 521	1 584	1 844	1 906	2 127	2 167	2 201	2 305	2 361
45-49	1 251	1 295	1 493	1 554	1 818	1 882	1 945	2 102	2 145

Source: Author's own calculations

Table 3.3 Yearly exponential growth rates

	1993-94	1998-99	2004-05*	2008-09
15-19	0.028	-0.011	-0.005	0.012
20-24	0.016	0.027	-0.010	-0.002
25-29	0.027	0.017	0.030	-0.006
30-34	0.025	0.029	0.014	0.035
35-39	0.036	0.019	0.027	0.019
40-44	0.041	0.033	0.016	0.024
45-49	0.035	0.040	0.033	0.020

**The growth rate for calculating the population at the median date of TDHS-2003 was based on the change between 2004 and 2005 because the median date of this survey falls in early 2004.*

Table 3.4 Populations at median survey dates (x1000)

	1993.69	1998.68	2004.06	2008.83
15-19	2873	3231	3058	2997
20-24	2665	2859	3209	3057
25-29	2366	2649	2876	3212
30-34	2178	2348	2672	2857
35-39	1879	2153	2345	2632
40-44	1533	1855	2152	2323
45-49	1259	1504	1855	2116

The obtained populations of five year age groups are then multiplied by the proportion of ever married women of the corresponding age group (${}_{x}m_{5}^{i}$), obtained from the household members' dataset of survey i.

$$ {}^{EM}_{x}N_{5}^{i} = {}_{x}N_{5}^{it} \times {}_{x}m_{5}^{i} $$

The proportions of ever married women from the household member data sets (${}_{x}m_{5}^{i}$) are given in Table 3.5, and the resulting populations (${}^{EM}_{x}N_{5}^{i}$) are given in Table 3.6.

Table 3.5 Proportions of ever-married women from the household members' datasets

	TDHS-1993	TDHS-1998	TDHS-2003	TDHS-2008
15-19	0.14	0.14	0.12	0.10
20-24	0.58	0.57	0.49	0.45
25-29	0.84	0.85	0.79	0.77
30-34	0.96	0.92	0.91	0.88
35-39	0.98	0.97	0.96	0.95
40-44	0.98	0.98	0.97	0.96
45-49	0.99	0.98	0.98	0.98

Table 3.6 Ever married women's populations at each survey date

	TDHS-1993	TDHS-1998	TDHS-2003	TDHS-2008
15-19	388,383	456,928	360,329	292,465
20-24	1,548,888	1,627,065	1,559,657	1,374,119
25-29	1,982,515	2,240,146	2,282,775	2,460,008
30-34	2,085,786	2,168,835	2,438,233	2,526,709
35-39	1,840,155	2,090,659	2,250,383	2,510,411
40-44	1,496,714	1,818,809	2,083,911	2,231,985
45-49	1,247,361	1,476,739	1,826,429	2,070,420
Total	10,589,803	11,879,182	12,801,717	13,466,118

Once the ever married population at survey date was obtained, the sample weights were inflated by a factor so that they would sum up to this population.

$$w_{ij}^{pop} = w_{ij} \times \frac{N_i^{EM}}{n_i}$$

As the last step, the population level weights were trimmed back down to add up to the total sample size of the four surveys.

$$w_{ij}^{adj} = w_{ij}^{pop} \times \frac{\sum_i n_i}{\sum_i N_i^{EM}}$$

The two adjustment stages presented above are demonstrated in Table 3.7

Table 3.7 Adjusting the sample weight to population weight and calibration to the original sample sizes

	Total ever-married population	Original weighted sample sizes	Population weight inflation factors	Calibrated inflation factors	Adjusted weighted sample sizes
TDHS-1993	10,589,803	6519	1624	0.940	6126
TDHS-1998	11,879,182	6196*	1917	1.109	6872
TDHS-2003	12,801,717	8075	1585	0.917	7406
TDHS-2008	13,466,118	7405	1819	1.052	7790
Total	48,736,820	28195			28195

**The number of weighted and unweighted women are the same for all surveys except for TDHS-1998; due to the presence of never-married women as well. The unweighted number is 6152.*

Table 3.8 Descriptives of the original and adjusted survey weights

		Sum	Mean	Standard deviation	Minimum	Maximum
TDHS-1993	Sample weight	6519	1.000	0.279	0.609	2.111
	Adjusted sample weight	5999	0.920	0.257	0.561	1.943
TDHS-1998 *Ever married*	Sample weight	6196	1.007	0.499	0.326	2.048
	Adjusted sample weight	6916	1.124	0.557	0.364	2.286
TDHS-1998 *Never married*	Sample weight	2380	0.982	0.486	0.326	2.048
	Adjusted sample weight	2656	1.096	0.542	0.364	2.286
TDHS-2003	Sample weight	8075	1.000	0.579	0.273	3.584
	Adjusted sample weight	7449	0.922	0.534	0.252	3.306
TDHS-2008	Sample weight	7315	1.001	0.579	0.248	2.911
	Adjusted sample weight	7735	1.058	0.534	0.262	3.079

As mentioned before, the sample for TDHS-1998 included never married women as well as ever-married women. The actual population of these women, however, was not used in the adjustment. The never-married women weight adjustment was made exactly the same as the ever-married women weights, so that the proportions of each within the sample would remain the same. With the adjustment, the weighted number of never-married women was changed from 2380 to 2656.

Other than the basic weight adjustment described in this section, a weighting scheme that gave equal weights to birth cohorts was used for the ideal number of children variable; and a weighting scheme that would adjust the number of women in a cohort to the age distribution of a specific survey was used for the mean number of children ever born variable. These are explained in Sections 3.4.3 and 3.4.4 below.

3.3. Sampling Errors

Sampling errors in sample surveys are usually estimated with special techniques. For a simple random sample, a stratified design with proportional allocation, or a design with equal cluster sizes, the calculations are straight forward. This is because simple statistics (in the form of means and proportions) are linear under these designs. As mentioned in the Weighting section, whenever designs are made in a way that selection probabilities are different among units, weighting becomes a requirement for correct statistical inference. With the introduction of weighting, all statistics become non-linear, with the numerator becoming a sum of weighted quantities, and the denominator becoming a sum of weights. The obtained statistic is a ratio estimator, and there are two approaches to obtain its sampling variance: (1) Model based methods (generally Taylor series linearization technique), (2) Resampling methods (jackknife repeated replication (JRR), balanced repeated replication, bootstrapping, etc.) (Oranje, 2006). Both approaches take stratification, clustering and weighting into account. This section explains how DHS computes sampling errors, and how they were computed from the combined dataset.

3.3.1. Sampling Errors in Turkey Demographic and Health Surveys

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 the variable that signifies 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 good spread among different geographical locations in the design strata, creating implicit strata; the PSUs in which have geographical proximity. DHS uses this property, and creates strata consisting of usually two PSUs each, like a paired selection design, and uses these strata when performing Taylor Linearization for variance estimation. Sampling error calculation is thus based on these combined strata, with usually 2 primary sampling units (PSUs, or here, clusters) each. Occasionally, there are three or more. This approach is usually preferred to simplify calculations.

Two of the datasets (TDHS-1993 and TDHS-1998) have variables that include information on these implicit strata (v022- Sample stratum number). Complex sample variances were calculated for means and proportions for the TDHS-1993 report, and rates were left out. It follows that the standard errors for the TFR and IMR cannot be observed in the TDHS-1993 final report.

No documentation is available on the formation of the v022 (Sample stratum number) variable in TDHS-1993 and TDHS-1998, thus they could not be replicated for this dissertation. Nevertheless, they have been used in this dissertation, so as to stick to the DHS approach and ensure comparability with the main report. For TDHS-2003 and TDHS-2008, implicit strata information was not found in the publicly available data sets, however, they were available in other formats, as used by DHS in the tabulations of the TDHS final reports. These variables were named “pairpsu”. Sampling errors of selected indicators were computed with these implicit strata (v022 for TDHS-1993 and TDHS-1998; pairpsu for TDHS-2003 and TDHS-2008) and checked with DHS final reports for consistency (Table 3.9)

The number of sampling error computation strata and sampling error computation units for the women’s data sets of TDHS-1993, TDHS-1998, TDHS-2003 and TDHS-2008 are 226-478, 235-476, 341-687 and 316-628 respectively.

Table 3.9 Stratification variable cross-check with TDHS reports

	Proportion	SE	CI lower	CI upper	DEFT	UW. number of cases
TDHS-1993						
Urban residence						
DHS	0,641	0,010	0,622	0,661	1,636	6519
SPSS (v022 as strata)	0,641	0,010	0,622	0,660	1,636	6519
Graduated secondary or higher						
DHS	0,175	0,008	0,159	0,191	1,664	6519
SPSS (v022 as strata)	0,175	0,008	0,160	0,191	1,664	6519
Currently married						
DHS	0,962	0,003	0,956	0,967	1,154	6519
SPSS (v022 as strata)	0,962	0,003	0,956	0,967	1,154	6519
Knows any method						
DHS	0,990	0,002	0,987	0,993	1,307	6519
SPSS (v022 as strata)	0,990	0,002	0,987	0,993	1,307	6519
Mean number of children ever born						
DHS	3,041	0,044	2,954	3,128	1,492	6519
SPSS (v022 as strata)	3,041	0,044	2,960	3,130	1,492	6519
TDHS-1998						
Urban residence						
DHS	0,665	0,017	0,631	0,699	3,299	8576
SPSS (v022 as strata)	0,665	0,017	0,631	0,697	3,299	8576
No education						
DHS	0,167	0,006	0,155	0,179	1,499	8576
SPSS (v022 as strata)	0,167	0,006	0,156	0,180	1,499	8576
With secondary education or higher						
DHS	0,303	0,010	0,282	0,324	2,107	8567
SPSS (v022 as strata)	0,303	0,010	0,283	0,324	2,107	8576
Currently married						
DHS	0,690	0,006	0,678	0,702	1,197	8567
SPSS (v022 as strata)	0,690	0,006	0,678	0,702	1,197	8576
Knows any method						
DHS	0,989	0,002	0,986	0,993	1,309	5893
SPSS (v022 as strata)	0,989	0,002	0,985	0,992	1,312	5893
Mean number of children ever born						
DHS	2,007	0,027	1,957	2,062	1,137	8576
SPSS (v022 as strata)	2,007	0,027	1,952	2,061	1,137	8576

Table 3.9 (continued). Stratification variable cross-check with TDHS reports

	Proportion	SE	CI lower	CI upper	DEFT	UW. number of cases
TDHS-2003						
Urban residence						
DHS	0,712	0,006	0,700	0,724	1,192	8075
SPSS (pairpsu as strata)	0,712	0,006	0,700	0,724	1,192	8075
No education						
DHS	0,218	0,008	0,202	0,234	1,777	8075
SPSS (pairpsu as strata)	0,218	0,008	0,202	0,235	1,777	8075
With secondary education or higher						
DHS	0,245	0,009	0,227	0,262	1,793	8075
SPSS (pairpsu as strata)	0,245	0,009	0,228	0,262	1,793	8075
Currently married						
DHS	0,950	0,003	0,945	0,955	1,121	8075
SPSS (pairpsu as strata)	0,950	0,003	0,944	0,955	1,121	8075
Knows any method						
DHS	0,998	0,000	0,997	0,999	0,757	7687
SPSS (pairpsu as strata)	0,998	0,000	0,997	0,998	0,756	7686
Ideal number of children						
DHS	2,510	0,020	2,469	2,551	1,183	4029
SPSS (pairpsu as strata)	2,509	0,020	2,469	2,549	1,180	4029
TDHS-2008						
Urban residence						
DHS	0,758	0,006	0,745	0,771	1,306	7405
SPSS (pairpsu as strata)	0,758	0,006	0,745	0,771	1,306	7405
No education						
DHS	0,183	0,008	0,167	0,200	1,796	7405
SPSS (pairpsu as strata)	0,183	0,008	0,168	0,200	1,796	7405
With secondary education or higher						
DHS	0,298	0,011	0,276	0,320	2,103	7405
SPSS (pairpsu as strata)	0,298	0,011	0,277	0,320	2,103	7405
Currently married						
DHS	0,945	0,004	0,937	0,953	1,463	7405
SPSS (pairpsu as strata)	0,945	0,004	0,937	0,952	1,463	7405
Knows any method						
DHS	0,998	0,000	0,997	0,999	0,880	7042
SPSS (pairpsu as strata)	0,998	0,000	0,997	0,999	0,877	7042

Table 3.9 (continued). Stratification variable cross-check with TDHS reports

	Proportion	SE	CI lower	CI upper	DEFT	UW. number of cases
Ideal number of children						
DHS	2,520	0,021	2,478	2,562	1,499	7261
SPSS (pairpsu as strata)	2,520	0,021	2,478	2,561	1,497	7261

The DHS approach in calculating sampling variances, thus standard errors, is via using Taylor series linearization for means and proportions. For more complex indicators such as demographic rates, DHS employs JRR. In statistics, Taylor series approximates reasonable estimates to functions, which are sometimes even linear functions (Lee and Forthofer, 2006). The Taylor series variance of a ratio mean x/z is approximated as:

$$var\left(\frac{x}{z}\right) \cong \frac{var(x) + R^2 \times var(z) - 2 \times R \times cov(x,z)}{z_0^2}$$

Where $R = \frac{x_0}{z_0}$ (which is the population value, best estimated by x/z).

The TDHS formulization and notation of Taylor series linearization for ratio statistic $r = y/x$ is as follows:

$$var(r) = \frac{1-f}{x^2} \sum_{h=1}^{H} \left[\frac{m_h}{m_h - 1} \left(\sum_{i=1}^{m_h} (y_{hi} - r x_{hi})^2 - \frac{(y_h - r x_h)^2}{m_h} \right) \right]$$

Where h denotes stratum, m denotes the number of clusters in stratum h, y denotes the weighted estimate, x denotes the weighted number of cases (i.e. sum of weights), and f denotes the finite population correction[2]; which is ignored due to the large sample sizes in DHS surveys.

---

[2] The finite population correction is used when the sample is a significantly big proportion of the population. It is calculated as $\left(1 - \frac{n}{N}\right)$. Whenever n is big, n/N approaches 1, reducing the sampling variance estimate. It follows that when the sample size approaches the population size, sampling error tends to 0; meaning there is no sampling.

Jackknife repeated replications technique was introduced by Quenouille in 1949. After further work by Tukey (1958) and Durbin (1959), Frankel (1971) used this method for variance estimation from complex sample surveys (Lee and Forthofer, 2006). The idea of jackknifing is resampling from the sample. One unit is removed, the desired statistic is calculated based on the remaining units, and then the initially removed unit is put back to remove another one, until all (but one) units have been removed once. The sampling variance is estimated by these statistics estimated from the replicates.

Replication methods as JRR have some advantages over Taylor series linearization. The major advantage is that they are useful for almost any type of estimate, including those that cannot be expressed in terms of formulas, such as the sample median or percentiles (Lee and Forthofer, 2006). However, Taylor series and JRR yield very similar results for simple statistics like the mean.

For stratified cluster designs, the units that are dropped from replicates are clusters. When a cluster is dropped from a stratum, the remaining clusters in the stratum are weighted in a way so that they would compensate for the dropped unit. If there are a units in the stratum, all clusters are multiplied by $(a-1)/a$ within that stratum. This is called a “Jackknife n” approach (Brick *et al.*, 2000). The formula for JRR n is as follows:

$$v(\bar{u}) = \sum_{h}^{L} \left(\frac{n_h - 1}{r_h} \right) \sum_{i}^{r_h} (u_{hi} - \bar{u})^2$$

Where h denotes stratum, L denotes the number of strata, r denotes replicate, u_{hi} is the replicate estimate, and $\bar{u}$ is the whole sample estimate (Lee and Forthofer, 2006).

The JRR employed in TDHSs is somehow different; it does not take stratification into account. Once a cluster is removed, the desired statistic is obtained from all remaining clusters, and no extra weighting is done for the clusters that share the same stratum as the dropped cluster. This approach is called a “Jackknife 1” approach (Westat, 2000). The TDHS sampling error JRR 1 formula for the ratio statistic $r = x/y$ is as follows:

$$var(r) = \frac{1}{k(k-1)} \sum_{i}^{k} \left[(kr - (k-1)r_{(i)}) - r \right]^2$$

Where k is the total number of clusters, $r_{(i)}$ is the replicate estimate, and r is the overall estimate.

DHS documentation (Macro International Inc. , 1996; ICF International, 2012) presents the JRR 1 formula and do not discuss why JRR 1 is preferred over JRR n, where, where JRR 1 is a special case of the general JRR n method (Westat, 2000). A potential reason could be that JRR 1 provides more liberal confidence intervals; which in turn could make policy makers / users take into account any potential sample bias, however, this suggestion does not go beyond the author’s reasoning.

3.3.2. Sampling Errors in the Combined Data Set

Sampling errors were calculated in taking stratification, clustering and weighting into account. A common stratification variable was created using the implicit strata variables from all data sets, as explained in the previous section. All strata were re-numbered, so that none of them would have the same stratification code. The clustering variable was also a new variable that would go from 1 to n, where n is the number of clusters in the sampling error computation strata. The weighting variable was obtained as explained in Section 3.2.2.

To keep a standard approach, the sampling variances of all selected statistics were obtained with Taylor series approximation or Jackknife repeated replications, depending on what was chosen by DHS. However, as opposed to the regular TDHS approach, stratification was taken into account. For simple statistics, standard errors were obtained directly from software (SAS) and for more complicated statistics, SAS Macros were prepared that would manually compute the JRR. The details of the standard error estimations and SAS macros are provided for each statistic in Section 2.4.

3.4. Selected Statistics

The Turkey Demographic and Health Surveys include a wide variety of variables. Various proportions and means are computed based on these variables. This thesis aimed to combine data from multiple surveys so that these statistics could be calculated with higher precision. Data could be combined on the basis of either age groups or cohorts. A combined statistic on women aged 15-49 from two time points would reflect the combination of two different 35-year birth cohorts of women with a 30 year overlap, observed 5 years apart. The choice of combining over cohorts or ages would be up to the interest of the researcher and the variables to be studied.

Calculating statistics on the basis of birth cohorts would yield estimates that are free from cohort effects, fixing the year of birth. For example, TDHS-1993 is based on a sample of women who were born through 1944-1978, and TDHS-1998 data is based on women born through 1949-1983. Both surveys include observations on the 1949-1953 birth cohort, and they could be combined to provide a statistic, the nature of which would not change over time. In other words, combining for a time-invariant variable over cohorts would serve to produce more precise point estimates with smaller standard errors.

However, recall errors or selection bias can arise in combining cohort estimates over time. If the time-invariant variable of interest refers to an information that relates to the younger ages of life (e.g. childhood place of residence, age at first marriage, educational level etc.), women who will be reporting in later surveys will be reporting about events that are more distant in the past. Thus recall errors would be increased for such women. Selection bias can arise due to mortality, or other causes that would exclude some women out of the sample (institutional population, emigration, etc.). Women with some characteristics that are related to the variable of interest may have higher rates of mortality, which will eventually bias the later estimates. For example, for a setting where maternal mortality is high, women who survive through later ages would be those that tend to have fewer births.

Combining time-variant variables on the basis of birth cohorts could also provide useful information, depending on the interest of the researcher. However, such a procedure would introduce age effects to the combined estimates. Such mean or a proportion could be higher or lower at a certain age than an earlier age. In this case, the time period to which the estimate belongs should be specified when presenting estimates. For example, one could combine the contraceptive prevalence of the 1949-1953 birth cohort from TDHS-1993 and TDHS-1998, which would show the average contraceptive prevalence for this cohort between 1993 and 1998. This is one of the applications taken up in this dissertation, and its details are found in Section 3.4.1.

3.4.1. Educational Attainment

The educational attainment variable was used as an example of combining proportions across cohorts over time. This variable is expected to remain constant for a given cohort across time; it is presumably time invariant. The reason is that most women get married upon completing a certain level of education. However, the educational level of younger women (especially the 15-19 age group) is not expected to remain constant overtime.

The educational status of any cohort of ever married women is expected to be stable over time, if 1) there is no educational differential regarding marital status and 2) mortality does not differ between different educational groups. The first condition affects the youngest age groups the most: As the marital composition of a cohort of ever married women changes, the mean level of education for this cohort should also change in the absence of this first condition. Previous studies from TDHS data showed that early marriage (Tezcan and Adalı, 2012) and childbearing (beginning only after marriage in Turkey) is more common among less educated women (Koç and Ünalan, 2000). This suggests that the 15-19 age group in the 1993-TDHS will be less educated on average than the 20-24 age group in the 1998-TDHS, although they are the same birth cohort.

The two conditions mentioned in the first sentence of the above paragraph are related to a concept known as **selection bias** (Winship and Mare, 1992), where observations are not dependent of the outcome (as the marital status of a cohort changes, the educational composition also changes). In the first case, women who are never married at survey date are not selected, and in the second case women who fail to survive from one survey to the next are not selected. If, for instance, women who are less educated are somehow more disadvantageous in terms of health, which in turns lead to higher mortality, then as one follows the educational status of a cohort in time, it should increase.

While it is possible to get an insight about the relationship between education and marital status by analyzing women of all marital status from the household members data set, the mortality bias remains a mystery, given there is no information on women who did not survive up to 50 years of age. Only women who survived until survey date are included in the sample.

All four datasets include education variables that were used in the final main reports for tabulations. However, the four variables are not standard. In the recoded data sets, the educational attainment variables are directly obtained from the different combinations of the following questions in the questionnaire: “Have you ever attended school?”, “What is the highest level of school you attended?”, “What is the highest grade you completed at that level?” and “Did you graduate from this school?”. Different combinations of these questions were used for the educational attainment variable in different datasets. Therefore a new variable with eight categories was created with the following categories: (1) Never attended any school (2) Attended primary school/education[3], but did not complete five years (3) Completed primary school or the first 5 years of primary education (4) Attended secondary school but did not complete it / attended the second level of primary education (6-8 years) but did not complete it (5) Graduated from secondary school or primary education (6) Attended high school but did not complete it (7) Completed high school (8) Attended university and higher. Questions on highest level of school attended, graduation from school, highest grade attended in school and whether it was completed were used, with help from the education in single years and

---

[3] The compulsory education system in Turkey consisted of a requirement of a primary school degree, which corresponded to 5 years of schooling until 1997. Secondary school was optional and had a duration of 3 years. This changed in 1997, and the new requirement was a total of 8 years of compulsory education, resulting in the abolishment of secondary schools. The new 8 year schooling was named primary education.

educational attainment variables. This variable was recoded into a three category variable and used in analysis.

Thus educational level was assessed in three categories: No education or first level primary education incomplete, first level primary education complete or second level primary school incomplete, and primary education complete or higher (including high school/university and postgraduate attendants or graduates). Since the estimates were calculated for cohorts rather than age groups, the weighting method described in Section 2.2.2 was used.

3.4.2. Contraceptive Use

As an application of combining statistics across age groups; a time variant variable; contraceptive use was calculated from the combined TDHS women data set. The contraceptive use variable is consistent across the three surveys and has the following categories: (1) No contraceptive use, (2) Folkloric contraceptive use, (3) Traditional contraceptive use and (4) Modern contraceptive use.

The data set was weighted by the adjusted weight described in Section 2.2.2. Since this weight corrects the sample sizes according to population trends, no further adjustments were made. The combination was made for women aged 15-49. The results were interpreted as the average proportions of women who do not use contraceptives, women who use traditional contraceptives and women who use modern contraceptives between 1993 and 2008. Ideal number of children was selected as a variable to exemplify the combination of time-variant variables across birth cohorts (Section 2.4.3), thus this variable was not computed for birth cohorts.

3.4.3. The Mean Number of Ideal Number of Children

The ideal number of children ever born is selected as an example of a combined mean across birth cohorts in time. The variable is obtained from the questionnaire with the following question: “If you could go back to the time you did not have any children and could choose exactly the number of children to have in your whole life, how many would that be?”. This hypothetical question results in a time-variant variable: Women may define a different number of ideal at different survey dates, thus this number can change over the life course for women: For example, older women may adapt their ideals to meet their achieved fertility (Testa, 2012).

The cohorts based estimates for this indicator is provided by giving equal weight to each cohort; so that it provides a simple average of different surveys. These weights were calculated in a way that their sum would remain the same. If, for example, a combined data estimate of TDHS-1993 and TDHS-1998 were to be obtained for cohort *c*, where the sizes of cohorts are n_c^{1993} and n_c^{1998}; the sample weight of each individual in cohort *c* in TDHS-1993 was multiplied by $\frac{n_c^{1993}+n_c^{1998}}{2\times n_c^{1993}}$. The ideal number of children of cohort *c* from these two surveys can be interpreted as the mean number of ideal number of children of cohort *c* between 1993 and 1998.

3.4.4. The Mean Number of Children Ever Born

The mean number of children ever born to all women aged 15-49 is one of the key variables presented in the fertility sections DHS reports. It is usually both calculated for ever married women, and all women in these reports. Here, the all women version was chosen an example of calculating a ratio mean from the combined women’s data set. The numerator for mean number of children ever born is the overall number of live births given by women weighted by the sample weight, and its denominator is the number of women, weighted by the sample weight and all women factors.

$$MCEB = \frac{\sum_j b_j w_j}{\sum_j w_j awf_j}$$

Where j is an index for women in the data set.

In the TDHS reports, the mean number of women is calculated for each woman based on the number of births she has given up to the survey date, based on multiple questions on total children ever born. The questions are the number of male/female (asked separately) children that live with the women, the number of male/female children that do not live with the woman, and the number of male/female babies they have given birth to who are not alive. These questions are consistent with the birth history sections in DHS surveys, both because of editing in the field, and consistency checks during data entry. Birth history data had to be used for the calculation of the mean number of children in this thesis, so that dates can be standardized and data can be combined across surveys.

The mean number of children for women is a cumulative, cohort measure that is time dependent. Thus there are age effects on the mean number of children ever born to a birth cohort of women from one survey to the other. A given birth cohort of women, at a given survey date, should have a mean number of children that is greater than or equal to that which was observed at the previous survey, unless there are serious sampling or non-sampling errors. This is because they are 5 years older, and have been exposed to childbearing during these five years.

In order to eliminate these age effects, the calculations are fixed for time points for five year birth cohorts as such:

$$MCEB_{c,t} = \frac{\sum_{i}^{n_c} b_{i,t} w_i}{\sum_{i}^{n_c} w_i awf_i}$$

Where $b_{i,c,t}$ is the number of birth by woman i in cohort c, given until time t.

The oldest cohort is born through 1944-1948, which roughly corresponds to the 45-49 age group in TDHS-1998. However, the first birth cohort for which we can have estimates from at least two surveys (TDHS-1993 and TDHS-1998) is the next birth cohort; 1949-1953. In order to make sure consecutive surveys cover the same period, we limited the birth history section of the more recent survey to the earliest month of interview of the older survey. For TDHS-1993, this is September, the month when the fieldwork started in. Limiting the birth history of TDHS-1998 up to September 1993, as if women were interviewed at that date, allows the improvement of the precision of estimates obtained from TDHS-1993. More data was added by limiting the birth history sections of TDHS-2003 and TDHS-2008 as well. Up to three data sets were used for the re-estimation of TDHS-1998 means by limiting the birth histories at September 1998, and up to two data sets for TDHS-2003, setting the limit as December 2003.

In addition to separate birth cohort estimates, the overall mean number of children was also re-estimated from multiple data sets for each survey date. The overall mean number of children ever born to women for a specific survey is implicitly weighted by the age structure observed in that survey. If one chooses to estimate this mean by the means for 5-year birth cohorts, the weighted average of the seven cohort means would provide this, where the weights are the number of all women in the birth cohorts.

$$MCEB = \frac{\sum_c n_c MCEB_c}{\sum_c n_c}$$

While there are mainly seven birth cohorts per survey, a survey conducted 5 years later can provide data on only six of these, rather than seven (the oldest birth cohort in the original survey is not interviewed in the newer survey). Thus the age distribution of the combined data gets distorted, leaving fewer cases in the oldest cohort. To overcome this problem, the cohort means were weighted by the cohort distribution of the original survey. If this adjustment is avoided, then data would be accumulated around younger ages, pulling the mean down.

Since the mean number of children has an extra weight in its denominator (all women factor), it is not a mean whose standard error can be calculated by the means procedures of computer software packages. Therefore it was computed through a SAS macro that used Jackknife repeated replications to compute its standard error.

$$SE^2(MCEB_c) = var(MCEB_c) = \sum_{\alpha=1}^{a} \sum_{\beta=1}^{b_\alpha} (MCEB_{c_{\alpha\beta}} - MCEB_c)^2$$

Where the number of clusters in stratum α is b_α, the total number of strata is a, CEB_c is the mean number of children ever born for cohort c, and $CEB_{c_{\alpha\beta}}$ is the mean number of children ever born to cohort c, computed on the replicate from strata α, where cluster β is deleted, and the remaining (b_α-1) elements in that strata are weighted up by the factor $b_\alpha/(b_\alpha - 1)$.

For the overall mean number of children ever born, additional steps were included in the macro. No matter how many surveys are included, the macro takes the age distribution of the first survey, and computes a weighted mean based on this distribution. This is also repeated at each step of the Jackknife Repeated Replications loop: Whenever a cluster is dropped, the new age distribution of the initial survey is re-computed, and used for the overall weighted mean number of children ever born.

The SAS macro is invoked by the following line:

%MACRO MCEB(DATA, STRATA, CASEID, DOMAIN, AWFDOM, SURVEY, INTDATE);

Where DATA is the dataset, STRATA is the variable that specifies the implicit strata, CASEID specifies cases (a new variable is required if one wishes to use more than one dataset), DOMAIN is the domain variable of interest, AWFDOM is the all women factor for the specified domain, SURVEY is the survey IDs which are included (filled as “%NRBQUOTE(1)” for TDHS-1993, and “%NRBQUOTE(1,2)” for TDHS-1993 and TDHS-1998 together; to include commas and the values for the surveys need to be predefined for this macro variable to work) and INTDATE is to determine the upper time limit of the birth history data (in century month codes). The SAS macro can be found in Appendix 1.

3.4.5. Total Fertility Rate

The total fertility rate, by definition, is the sum of age-specific fertility rates, which are based on the number of births classified by mothers’ age and the number of women who have given these births classified by age group during a specific period of time. It is a period measure, and is usually calculated from data on women aged 15-49 (sometimes 15-44), regardless of marital status (all women).

Age-specific fertility rates in Demographic and Health Surveys are calculated for 5-year age groups as follows:

$$ASFR(i,t) = \frac{b(i,t)}{e(i,t)}$$

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* (Rutstein, 2002).

The date of birth of the child and age of mother at this birth are checked for each birth of the woman for the numerator, before assigning them as births to age group *i*, that covers the ages *[a,b]*:

$$b(i,t) = \sum_j w_j I_{[a,b]}$$

where

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

and w_j is the sample weight for woman *j*.

For the denominator, there are several checks to be made. During the time period *t*, a woman might be exposed to childbearing within only one 5-year age group, multiple 5-year age groups, or might be out of the reproductive age range for a part of the period. If the woman's age at the beginning of period is age_{min}, and her age at the end of the period is age_{max}, then the exposure to age group *i* during time period *t* is:

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

Where

$$EXP_{[a,b]} = \begin{cases} [b - \max(a, age_{\min})] & \text{if } age_{min} < b < age_{max} \\ [age_{max} - \max(a, age_{\min})] & \text{if } a < age_{max} < b \end{cases}$$

and w_j is the sample weight.

The total fertility rate by definition is based on data from all women. As mentioned earlier, since interviews were carried out with only ever-married women in TDHS-1993, TDHS-2003 and TDHS-2008, all women factors have been used to inflate the denominators of the TFR to account for never-married women as well. The factors were set to 1 for TDHS-1998, where all women regardless of marital status were interviewed. The all women factor operates similar to the sample weight during calculation; the exposure value of each case is multiplied by this factor to change the above formula to

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

where awf_j is the all women factor for woman j. This factor takes separate values for each single age, and is calculated separately for the levels of variables that exist for both ever married and never married women in the data set (e.g. urban-rural, regions). For a detailed description of the calculation of the all women factors, see Rutstein (2003).

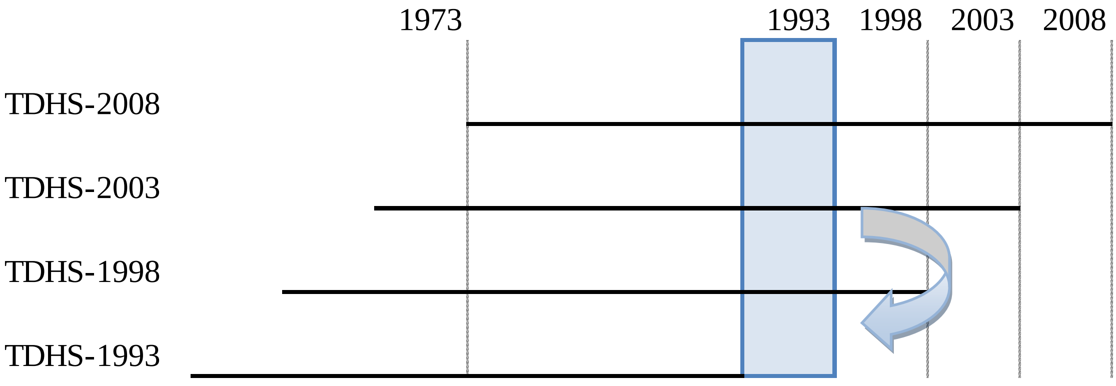
Since Demographic and Health Surveys collect data on births through retrospective birth history sections, data is available from the beginning of women’s reproductive period up to the survey date. When calculating the TFR from multiple DHSs, women who are interviewed in different surveys but are born during the same calendar period constitute the new age groups. The 15-19 age group in one survey, becomes the 20-24 age group in the next survey that takes place 5 years later.

For example, a woman born in 1984 and interviewed for TDHS-2008 is 24 years old at survey date. She would be observed for 9 years of her reproductive ages, assuming this period begins at the age of 15. Data on this woman can be used to calculate the TFR, as if she were interviewed in the previous survey, TDHS-2003. Back at this date, she would be 19 years old, having been exposed to the 15-19 age group for 3 years.

Thus, such cases can be included in the calculations of TFR from both surveys, because both births and exposures are known up to each survey date. Total fertility rates published in DHS reports go back a fixed amount of time (usually 1, 3 or 5 years) starting from survey date. It is possible to pull this date backwards in time, for instance from September 2008 to September 2003, to use TDHS-2008 observations for the calculation of a 3-year TFR from 2008 and backwards. Figure 3.1 illustrates the idea with an example: While one can use only one single dataset to compute a TFR for any period after 2003 (TDHS-2008 data only), it is possible to use two datasets for the period after 1998 and before 2004 (TDHS-2003 and TDHS-2008). The earliest date for which all four datasets can be employed is 1993.

There are several issues regarding this procedure. One is attrition: There are women from seven different age groups in TDHS-2008 (from 15-19 to 45-49) when the TFR is calculated for a period after 2003. However, if the information for the same women is used for the calculation of a rate that goes back earlier than 2003, women in the younger age groups in the recent survey start getting lost due to this shifting back. For example, women who are aged 15-19 in TDHS-2008 are the same birth cohort of women as children who are aged 10-14 in TDHS-2003, thus they no longer provide data. Similarly, the oldest women interviewed in TDHS-2008 are 45-49 year olds, and they are aged 40-44 at the TDHS-2003 survey date. Thus the latter survey provides no data on 45-49 age group when a TFR for 2003 is calculated; TDHS-2008 yields in 6 age groups when taken five years back, rather than 7. This is barely a problem for this particular example, because the age-specific fertility rate for the 45-49 age groups is extremely low in Turkey. However, as surveys are taken further back in time, it becomes an issue.

Figure 3.1 Graphical representation of pooled data for the calculation of TFR



Data for a TFR based on all four data sets

Thiam and Aliaga (2001) developed a SAS Macro to calculate the TFR and its standard error. To calculate the number of births to an age group of women, this Macro creates a temporary child dataset, and creates indicator variables to determine whether these births will be included in the calculation, and if so, to which age group. The indicator variables are weighted and summed up to obtain the numerators of the age-specific fertility rates. Age-specific exposures are calculated based on women’s ages at the beginning and end of the interval for which the TFR is calculated. They are then weighted by the sample weight and all women factors, to provide the denominators of the age-specific fertility rates. The total fertility rate is the sum of the ASFRs multiplied by 5 (since the length of age groups is 5).

Since TFR is the sum of 7 ratios, the calculation of its standard error is not straight forward. Variance estimation techniques such as Taylor Series approximation or replication methods are suitable for such statistics. For rates calculated for Turkey; DHS prefers Jackknife repeated replications, based on cluster level weighted replicates. According to the final reports of the surveys (HUIPS; 1999, 2004, 2009), stratification was not taken into account at this stage, a Jackknife 1 approach was used. The SAS Macro by Thiam and Aliaga (2001) use the same approach for the standard error calculation of the TFR. It calculates the Jackknife variance with a loop that goes from 1 to the total number of clusters in the sample, where one cluster is dropped each time, and the replicate TFR is computed from the remaining clusters. The formula used by DHS in TDHS-1998, TDHS-2003, TDHS-2008 and the SAS Macro for the standard error of the TFR is as follows:

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

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 TFR) 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) (Thiam and Aliaga, 2001) .

The Macro written by Thiam and Aliaga (2001) is based on 6 parameters, and runs with the given line to invoke the TFR macro:

%TFR(DATA=,PSU=,YEARS,=BYVAR=,WEIGHT=,INFLATE=)

Where DATA is the individual level dataset to be used, PSU is the cluster variable, YEARS indicates the length of period for which TFR is to be calculated (prior to survey date), BYVAR specifies the domain variable, WEIGHT is the sample weight variable, and INFLATE is the all women factor variable[4]. Once the TFR macro is invoked, it produces an output for total and specified domains, which includes the estimate, its standard error, relative error, the weighted case numbers and the lower and upper confidence limits of the estimate.

This received SAS macro was modified to suit the needs of this study. First of all, a hard coded variable was added for the total all women factor. The domain specific all women factor was included as a macro parameter, given different domains might be requested when running the macro.

An estimation procedure for the standard error under a simple random sampling assumption was also added to the code. We used Jackknife repeated replications to estimate this statistic; however, it operates on individual level, rather than cluster level (similar to the Jackknife 1 approach, where cases are deleted one at a time rather than clusters of cases).

The formula used in this thesis for the estimation of the standard error under an SRS assumption is as follows:

$$SE^2(TFR) = var(TFR) = \sum_{j=1}^{n} (r_{(j)} - r)^2$$

---

[4] M. Thiam has kindly provided me with the SAS macros they have used. The INFLATE parameter was not included in the provided macro despite its appearance in the macro invoking line. It was a macro for surveys where all women are interviewed, such as TDHS-1998.

where

n is the number of women included in the dataset,

r is the TFR based on all cases,

$r_{(j)}$ is the replicate TFR (based on n-1 women, excluding the j^{th} woman).

The standard error estimation under the existing survey design was also modified. The reason for this is that all four TDHSs are weighted, stratified cluster designs, yet the standard errors provided in Turkey DHS reports do not take stratification into account; they are based on clusters only (Jackknife 1 approach, as explained in Section 3,3). On the other hand, this is not the case for means and proportions, stratification is accounted for when the Taylor series approximation is used for the standard error estimations. The strata used in these calculations are not the design strata, but are implicit strata, that have usually two or sometimes more clusters each. The same strata variables were used for the Jackknife replicates. When forming the replicates under stratification and clustering; one cluster is dropped from a stratum, and the remaining clusters in the strata are weighted up to compensate for the deleted cluster (Jackknife n approach). The number of replicates is the total number of strata subtracted from the number of clusters. If the number of clusters in stratum α is b_α, and the total number of strata is a, the formula for the standard error of the TFR is as follows:

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

where

r is the overall total fertility rate without any deletion,

$r_{\alpha\beta}$ is the TFR computed on the replicate from strata α, where cluster β is deleted, and the remaining (b_α-1) elements are weighted up by the factor $b_\alpha/(b_\alpha - 1)$.

Another addition we made to the original macro is the estimation of standard errors and design effects of the age-specific fertility rates, which are normally not presented within DHS reports in Turkey. The idea for calculating the standard error and the standard error under simple random sampling is the same as that of TFR's; Jackknife repeated replications was used for both.

As a result of the above modifications, the macro is invoked with the line below:

%TFR(DATA, INTDATE, MONTHS, STRATA, CASEID, DOMAIN, AWFDOM, SURVEY)

Where DATA is the dataset, INTDATE is the end of the period for which TFR will be calculated (in century month codes), MONTHS is the length of period, STRATA is the variable that specifies the implicit strata, CASEID specifies cases (a new variable is required if one wishes to use more than one dataset), DOMAIN is the domain variable of interest, AWFDOM is the all women factor for the specified domain, and SURVEY is the survey IDs which are included (filled as "%NRBQUOTE(1)" for TDHS-1993, and "%NRBQUOTE(1,2)" for TDHS-1993 and TDHS-1998 together; to include commas and the values for the surveys need to be predefined for this macro variable to work).

The output of the above macro includes the standard error of the TFR and ASFRs under the design and under an SRS assumption; for both the overall and the selected domains, under selected surveys. The SAS macro used for the computation if TFR and its standard error is provided in Appendix 2.

3.4.6. The Infant Mortality Rate

Infant mortality rate shows the number of children who did not survive until the end of their first year of life per 1000 children. The birth history sections of demographic and health surveys allow the direct estimation of this rate. There are three possible approaches to the direct estimation of infant mortality and similar rates (Rutstein and Rojas, 2003). One is a vital registration approach, where data on the number of births and number of infant deaths from a specific time period is used; and separation factors are used to account for possible changes in the number of births in time. A second approach is a true cohort life table approach, resulting in true probabilities of death to birth cohorts of children. The third approach is a synthetic cohort life table approach that combines the experiences of different cohorts of children to provide a more recent rate, this is the approach used by DHS (Rutstein and Rojas, 2003).

The DHS uses the following age groups (in months) to compute the probabilities of death: 0, 1-2, 3-5 and 6-11. These individual death probabilities are subtracted from 1 to calculate survival probabilities. The product of the survival probabilities subtracted from one gives the probability of death within the first twelve months of life, namely the IMR:

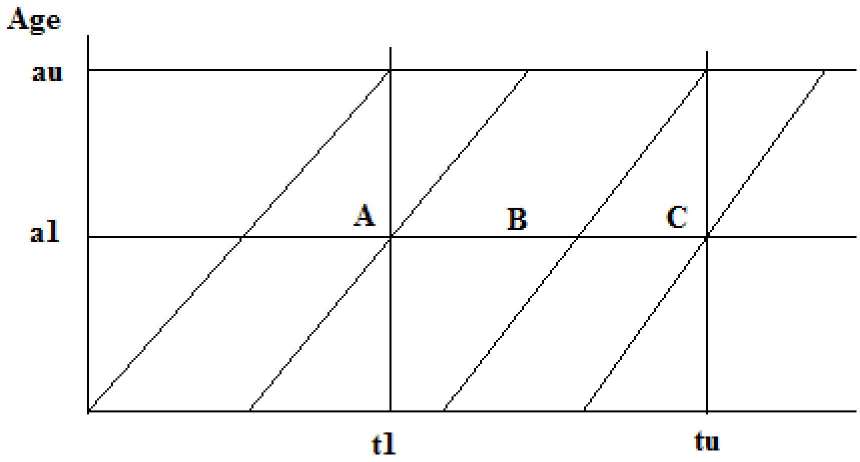
$$(n)q(x) = 1 - \coprod_{i=x}^{i=x+n} (1 - q(i))$$

Where x is the beginning of the age interval, and n is the length of the interval, estimating the probability of death between age x and $x+n$ (Rutstein and Rojas, 2003).

Since the infant mortality rate is based on synthetic cohorts in the DHS approach, all probabilities are defined by a time period, and an age interval. In Figure 3.1, the lower limit of the time period is denoted by t_1, the upper limit by t_u, the upper limit of the age interval by a_u, and the lower limit of the age interval by a_1. Three different possibilities arise: A child may be born prior to the period of interest, but may be observed within the age limit of interest during this period (cohort A). A child might be born and fully observed in the period of interest (cohort B). Finally, a child may be born towards the end of the period of interested, and may only observed partially for the age group of interest (cohort C).

While the estimation of the probability from cohort B in Figure 3.1 is straight forward, some assumptions are made for the other two cohorts. The main assumption is to assume that on average, these cohorts are observed for half the period of interest. There is an exception for the rates that end with the interview date: All deaths (rather than half) to cohort C are added to the numerator, because it is assumed to be already half of the deaths to cohort C between the age limits a_1 and a_u (Rutstein and Rojas, 2003), with the exception mentioned above.

Figure 3.2 Graphical representation of the calculation of birth probabilities based on a synthetic cohort approach



Source: Rutstein and Rojas, 2003

The way DHS takes these issues into account is through weighting. Possible combinations of date of birth, age limits and time periods are checked; while those observed for the whole period are weighted by the sample weight (cohort B), others are weighted by half the sample weight (cohorts A and C).

The idea of calculating IMR from multiple datasets is very similar to that of TFR (see Section 3.4.5), yet somehow more straight forward. The major difference is that the datasets that are combined are birth history datasets, rather than women's data sets. Each survey includes approximately a 35-year birth history section, 30 years of which overlap with the survey preceding it, almost doubling the sample size for the overlapping period.

Infant mortality rates are usually presented based on the 5 years preceding interview date in Turkey. By combining data across four DHS surveys, the rates for TDHS-1993, TDHS-1998 and TDHS-2003 can be re-calculated with more cases.

Thiam and Aliaga (2001) have developed a SAS Macro to calculate childhood mortality rates (neonatal mortality rate, postneonatal mortality rate, infant mortality rate, child mortality rate and under five mortality rate) with their standard errors. This macro first tabulates deaths, and then it tabulates exposures with a separate macro. The exposure macro runs twice, once with the sample weights, and once with every child getting a weight of 1; so that unweighted number of cases is also tabulated.

Just like the TFR, DHS uses Jackknife repeated replications to estimate the sampling error of the IMR for Turkey. It is a cluster level Jackknife and stratification is not accounted for. The Jackknife formula used by DHS is the same as that given in Section 3.31.

Unlike the TFR macro, the IMR macro estimates the sampling variance under a simple random sampling approach as well, thus estimates of the design factors are also presented. The way this sampling variance is estimated is via treating the rates as proportions, and the sampling variance takes the below form under this assumption:

$$var_{SRS}(r) = \frac{r(1-r)}{n}$$

where r is the rate of interest. The denominator n is the unweighted number of cases.

The IMR macro by Thiam and Aliaga (2001) is invoked with a line similar to the TFR, as below:

%IMR(DATA=,PSU=,YEARS,=BYVAR=,WEIGHT=)

Where DATA is the individual level dataset to be used, PSU is the cluster variable, YEARS indicates the length of period for which TFR is to be calculated (prior to survey date), BYVAR specifies the domain variable and WEIGHT is the sample weight variable. Once the IMR macro is invoked, it produces an output for total and specified domains, which includes the estimate, its standard error, relative error, standard error under a simple random sampling assumption, the design factor, the weighted and unweighted case numbers and the lower and upper confidence limits of the estimate.

As with the TFR macro, some modifications were introduced to the IMR macro. The way the SRS standard errors was estimated was changed to be a case-wise Jackknife repeated replications estimation, using the same idea as the TFR standard error under simple random sampling.

The sampling variance approach was also changed from Jackknife 1 to Jackknife n. The formula used is the same as TFR, given in Section 3.4.6.

After the modifications, the new IMR macro is invoked by the following line:

%IMR(DATA, PSU, BYVAR, WEIGHT, INTDATE, SURVEY)

Where DATA is the dataset, PSU is the strata code, BYVAR is the domain variable of interest, WEIGHT is the sample weight, INTDATE is the end of the period of interest, and SURVEY is the survey IDs to be included when running the code (filled as “%NRBQUOTE(1)” for TDHS-1993, and “%NRBQUOTE(1,2)” for TDHS-1993 and TDHS-1998 together to make commas work; and the values for the surveys need to be predefined for this macro variable to work).

The output of the above macro includes the five different childhood mortality rates, their standard errors under the given complex sample design and under an SRS assumption; for both the overall and the selected domains, under selected surveys. The SAS macro used for the computation if early age mortality rates and their standard errors is provided in Appendix 3.

4. RESULTS

The aim of this thesis was to combine the datasets of the last four demographic and health surveys in Turkey as an attempt to provide more precise statistics compared to those obtained from single datasets. The combined DHS dataset was tested on various statistics. These statistics include means, proportions and rates presented in the DHS main reports.

Educational attainment and contraceptive use were selected as proportions, mean number of children ever born was selected as a mean; total fertility rate and infant mortality rate were selected as examples of rates.

4.1. Combined Data Estimates of Proportions and Means

4.1.1. Proportions: Educational Attainment and Contraceptive Use

Educational Attainment

The first variable is a proportion, educational attainment of women. This variable exists in all four datasets, however, its definition is different. Therefore, all four were re-defined based on different variables, some of which are found in the non-recoded datasets (i.e. the previous versions of the finalized datasets). The final variable was created with eight levels: 1) Never attended school 2) Attended but did not graduate from first level primary school 3) Graduated from first level primary school, never attended secondary level primary school 4) Attended but did not graduate from secondary level primary school, 5) Graduated from secondary level primary school, never attended high school, 6) Attended but did not graduate from high school, 7) Graduated from high school, never attended university. 8) Any educational level equal to or higher than university attendance. This eight level variable was recoded to have three levels in order to obtain a balanced number of cases throughout the four surveys. The 1st and 2nd categories were combined, the 3rd and 4th were combined, and the remaining four categories (5, 6, 7 and 8) were combined.

As discussed in Section 3.4.1, the educational status of any cohort of ever married women is expected to be stable over time, especially for cohorts the marital composition changed in time. Therefore when presenting results, the youngest age group (15-19) was excluded, assuming the educational level of later age groups would change less in the later years, provided the 20-25 age group is where marriage peaks in Turkey for TDHS-1993, TDHS-1998 and TDHS-2003 (HUIPS 1994; 1999; 2004).

Table 4.1 shows the results for the education variable from all surveys separately, and then from the combinations of two, three and four surveys. The results are fairly consistent for the older age groups of all surveys. However, for the younger age groups, inconsistencies are observed just as expected, with confidence intervals that are either overlapping, or not overlapping at all. For example, for the 1979-1983 cohort, the proportion of women who fell in the lowest educational category was estimated as 27 percent alone from TDHS-2003, and it decreased to 19 percent with the inclusion of TDHS-2008; this shows the addition of data on more educated women who married later.

The patterns seen for different education categories are different. When a difference is observed in the first or second category, it is a downward trend towards the later survey. In other words, the proportion of less educated women among a given cohort seems to decrease over time. On the other hand, the corresponding proportion for the highest education category rises over time. The most likely explanation is the one discussed above, that there is a correlation between early marriage and low education, so the educational level of a cohort improves over time as more educated women get married and get included in the ever married cohort. One potential reason is the selection bias arising from mortality, again, as discussed above.

Table 4.1 Educational attainment of 5 year birth cohorts and its standard error from cumulations of TDHS-1993, TDHS-1998, TDHS-2003 and TDHS-2008

		1993				1993&1998				1993,1998&2003				All 4			
		Percent	SE	DEFF	CV	Percent	SE	DEFF	CV	Percent	SE	DEFF	CV	Percent	SE	DEFF	CV
1944-48	No edu/pri inc	56.77	2.39	1.480	0.04	56.70	2.37	1.46	0.04	56.70	2.37	1.46	0.04	56.70	2.37	1.46	0.04
	Pri comp/sec inc	32.96	2.01	1.152	0.06	33.17	2.00	1.15	0.06	33.17	2.00	1.15	0.06	33.17	2.00	1.15	0.06
	Sec comp and higher	10.27	1.29	1.116	0.13	10.14	1.27	1.12	0.12	10.14	1.27	1.12	0.12	10.14	1.27	1.12	0.12
1949-53	No edu/pri inc	50.33	1.92	1.220	0.04	49.90	1.46	1.14	0.03	49.90	1.46	1.14	0.03	49.90	1.46	1.14	0.03
	Pri comp/sec inc	38.78	1.68	0.987	0.04	38.41	1.32	0.97	0.03	38.41	1.32	0.97	0.03	38.41	1.32	0.97	0.03
	Sec comp and higher	10.89	1.37	1.629	0.13	11.69	1.09	1.41	0.09	11.69	1.09	1.41	0.09	11.69	1.09	1.41	0.09
1954-58	No edu/pri inc	39.40	1.63	1.127	0.04	40.51	1.24	1.07	0.03	38.15	1.01	1.04	0.03	38.14	1.00	1.04	0.03
	Pri comp/sec inc	47.63	1.71	1.173	0.04	44.92	1.22	0.99	0.03	46.50	1.04	1.03	0.02	46.52	1.04	1.03	0.02
	Sec comp and higher	12.98	1.32	1.544	0.10	14.57	1.02	1.29	0.07	15.35	0.87	1.33	0.06	15.34	0.87	1.33	0.06
1959-63	No edu/pri inc	27.65	1.65	1.608	0.06	28.39	1.13	1.28	0.04	28.70	0.90	1.21	0.03	28.49	0.79	1.16	0.03
	Pri comp/sec inc	54.78	1.73	1.456	0.03	53.36	1.20	1.17	0.02	52.37	1.04	1.29	0.02	52.29	0.92	1.22	0.02
	Sec comp and higher	17.57	1.29	1.401	0.07	18.25	1.09	1.55	0.06	18.93	0.95	1.68	0.05	19.22	0.87	1.65	0.05
1964-68	No edu/pri inc	23.84	1.58	1.542	0.07	21.78	1.02	1.28	0.05	22.22	0.81	1.26	0.04	22.39	0.73	1.28	0.03
	Pri comp/sec inc	56.55	1.74	1.407	0.03	57.07	1.29	1.33	0.02	56.61	1.03	1.32	0.02	56.30	0.91	1.28	0.02
	Sec comp and higher	19.60	1.39	1.384	0.07	21.15	1.19	1.61	0.06	21.17	0.92	1.51	0.04	21.32	0.81	1.38	0.04
1969-73	No edu/pri inc	20.95	1.70	1.517	0.08	17.85	1.00	1.37	0.06	18.21	0.82	1.43	0.04	18.59	0.72	1.42	0.04
	Pri comp/sec inc	61.31	1.72	1.147	0.03	60.94	1.27	1.26	0.02	58.93	1.01	1.24	0.02	58.52	0.89	1.25	0.02
	Sec comp and higher	17.75	1.34	1.169	0.08	21.21	1.11	1.29	0.05	22.85	0.87	1.23	0.04	22.89	0.81	1.37	0.04
1974-78	No edu/pri inc	19.34	2.48	0.997	0.13	15.23	1.16	1.15	0.08	14.51	0.86	1.49	0.06	13.77	0.68	1.45	0.05
	Pri comp/sec inc	71.89	2.79	1.035	0.04	64.95	1.54	1.05	0.02	59.44	1.26	1.43	0.02	57.15	1.07	1.39	0.02
	Sec comp and higher	8.76	1.66	1.050	0.19	19.82	1.25	0.93	0.06	26.05	1.09	1.27	0.04	29.08	0.99	1.32	0.03

Table 4.1 (continued). Educational attainment of 5 year birth cohorts and its standard error from cumulations of TDHS-1993, TDHS-1998, TDHS-2003 and TDHS-2008

		1993				1993&1998				1993,1998&2003				All 4			
		Percent	SE	DEFF	CV	Percent	SE	DEFF	CV	Percent	SE	DEFF	CV	Percent	SE	DEFF	CV
1979-83	No edu/pri inc					19.88	2.91	1.24	0.15	15.29	1.30	1.58	0.08	13.56	0.88	1.64	0.06
	Pri comp/sec inc					66.48	3.22	1.04	0.05	58.83	1.71	1.27	0.03	54.72	1.29	1.29	0.02
	Sec comp and higher					13.64	2.40	1.10	0.18	25.88	1.38	1.03	0.05	31.72	1.24	1.30	0.04
1984-88	No edu/pri inc									26.37	3.00	1.14	0.11	18.87	1.48	1.35	0.08
	Pri comp/sec inc									50.47	3.28	0.87	0.07	46.34	1.79	0.99	0.04
	Sec comp and higher									23.16	3.05	0.95	0.13	34.79	1.74	1.02	0.05
1989-93	No edu/pri inc													21.05	3.76	1.37	0.18
	Pri comp/sec inc													18.55	3.45	1.07	0.19
	Sec comp and higher													60.41	4.65	1.28	0.08

Note: Less than 25 cases were added to the 1944-1948 birth cohort from TDHS-1998.

Table 4.2 Educational attainment of 5 year birth cohorts and its standard error from cumulations of TDHS-1998, TDHS-2003 and TDHS-2008

		1998				1998&2003				1998&2003&2008			
		Percent	SE	DEFF	CV	Percent	SE	DEFF	CV	Percent	SE	DEFF	CV
1954-58	No edu/pri inc	41.63	1.86	1.02	0.04	37.51	1.27	1.01	0.03	37.50	1.27	1.01	0.03
	Pri comp/sec inc	42.17	1.73	0.86	0.04	45.91	1.31	0.98	0.03	45.95	1.31	0.98	0.03
	Sec comp and higher	16.20	1.56	1.16	0.10	16.58	1.14	1.27	0.07	16.56	1.13	1.26	0.07
1959-63	No edu/pri inc	29.12	1.54	1.05	0.05	29.22	1.08	1.07	0.04	28.76	0.90	1.06	0.03
	Pri comp/sec inc	51.93	1.67	0.96	0.03	51.18	1.30	1.23	0.03	51.47	1.09	1.17	0.02
	Sec comp and higher	18.94	1.75	1.65	0.09	19.61	1.26	1.77	0.06	19.77	1.08	1.70	0.05
1964-68	No edu/pri inc	19.98	1.33	1.08	0.07	21.53	0.94	1.15	0.04	21.96	0.82	1.21	0.04
	Pri comp/sec inc	57.52	1.87	1.28	0.03	56.63	1.26	1.29	0.02	56.22	1.07	1.26	0.02
	Sec comp and higher	22.51	1.87	1.74	0.08	21.85	1.17	1.55	0.05	21.81	0.97	1.39	0.04
1969-73	No edu/pri inc	15.68	1.19	1.22	0.08	17.25	0.92	1.39	0.05	18.05	0.78	1.39	0.04
	Pri comp/sec inc	60.68	1.80	1.32	0.03	58.09	1.22	1.27	0.02	57.88	1.03	1.27	0.02
	Sec comp and higher	23.64	1.62	1.33	0.07	24.65	1.08	1.24	0.04	24.07	0.95	1.39	0.04
1974-78	No edu/pri inc	14.10	1.31	1.21	0.09	13.94	0.92	1.57	0.07	13.36	0.71	1.50	0.05
	Pri comp/sec inc	63.05	1.82	1.06	0.03	57.98	1.37	1.47	0.02	56.07	1.12	1.41	0.02
	Sec comp and higher	22.85	1.53	0.95	0.07	28.08	1.19	1.29	0.04	30.56	1.05	1.33	0.03
1979-83	No edu/pri inc	19.88	2.91	1.24	0.15	15.29	1.30	1.58	0.08	13.56	0.88	1.64	0.06
	Pri comp/sec inc	66.48	3.22	1.04	0.05	58.83	1.71	1.27	0.03	54.72	1.29	1.29	0.02
	Sec comp and higher	13.64	2.40	1.10	0.18	25.88	1.38	1.03	0.05	31.72	1.24	1.30	0.04
1984-88	No edu/pri inc					26.37	3.00	1.14	0.11	18.87	1.48	1.35	0.08
	Pri comp/sec inc					50.47	3.28	0.87	0.07	46.34	1.79	0.99	0.04
	Sec comp and higher					23.16	3.05	0.95	0.13	34.79	1.74	1.02	0.05
1988-93	No edu/pri inc									21.05	3.76	1.37	0.18
	Pri comp/sec inc									18.55	3.45	1.07	0.19
	Sec comp and higher									60.41	4.65	1.28	0.08

In terms of the precision of estimate, the combined data estimates have notably lower standard errors. The highest decrease is observed with the addition of one survey to the original estimate to get a combined estimate from two datasets (Table 4.4). Among educational categories, the highest decrease in standard errors is observed for the lowest education category. The precision benefits of additional datasets decrease with more surveys added, reaching as high as 60% in total for some statistics with a total of 4 surveys (not shown). The coefficient of variation provides a clear display of the standard error lowering effect of adding more surveys. The design effects usually showed a downward trend for the relatively consistent estimates from older cohorts. For the 1979-1983 cohort where estimates differed more, the design effect increased.

Table 4.2 shows the combined data results for TDHS-1998 and later surveys. The findings are parallel to those observed in Table 4.1: Less consistent estimates for younger birth cohorts and decreasing standard errors with the addition of more surveys. Changes in design effects are less regular compared to Table 4.1. It increases for the older cohorts and younger cohorts, and changes by very little for those in between. The coefficients of variation demonstrate the decrease in standard errors, decreasing with additional surveys.

The findings from adding the TDHS-2008 results to TDHS-2003 results provide very consistent estimates of educational attainment among birth cohorts. The 1989-1983 birth cohort, being the 15-19 age group in TDHS-2003 has a higher educational level in the TDHS-2008, when they had become the 20-24 age group. The design effect increased for this birth cohort, and decreased for the rest.

Table 4.3 Educational attainment of 5 year birth cohorts and its standard error from cumulations of TDHS-2003 and TDHS-2008

		2003				2003&2008			
		Percent	SE	DEFF	CV	Percent	SE	DEFF	CV
1959-63	No edu/pri inc	33.22	1.69	0.98	0.05	33.21	1.69	0.98	0.05
	Pri comp/sec inc	49.80	1.97	1.12	0.04	49.86	1.96	1.12	0.04
	Sec comp and higher	16.97	1.65	1.38	0.10	16.94	1.65	1.39	0.10
1964-68	No edu/pri inc	29.31	1.52	1.09	0.05	28.59	1.12	1.06	0.04
	Pri comp/sec inc	50.45	1.99	1.50	0.04	51.24	1.40	1.27	0.03
	Sec comp and higher	20.24	1.82	1.90	0.09	20.18	1.36	1.73	0.07
1969-73	No edu/pri inc	23.01	1.34	1.24	0.06	22.95	1.04	1.28	0.05
	Pri comp/sec inc	55.78	1.70	1.29	0.03	55.58	1.30	1.25	0.02
	Sec comp and higher	21.22	1.43	1.32	0.07	21.47	1.11	1.21	0.05
1974-78	No edu/pri inc	18.84	1.41	1.53	0.07	19.21	1.01	1.45	0.05
	Pri comp/sec inc	55.48	1.65	1.22	0.03	56.50	1.25	1.25	0.02
	Sec comp and higher	25.68	1.42	1.14	0.06	24.29	1.18	1.43	0.05
1979-83	No edu/pri inc	13.82	1.27	1.92	0.09	13.09	0.84	1.64	0.06
	Pri comp/sec inc	54.20	1.94	1.76	0.04	53.53	1.37	1.52	0.03
	Sec comp and higher	31.98	1.71	1.51	0.05	33.37	1.30	1.44	0.04
1984-88	No edu/pri inc	13.98	1.45	1.75	0.10	12.82	0.92	1.73	0.07
	Pri comp/sec inc	56.65	1.97	1.32	0.03	53.34	1.39	1.32	0.03
	Sec comp and higher	29.38	1.59	0.99	0.05	33.84	1.34	1.32	0.04
1988-93	No edu/pri inc	26.37	3.00	1.14	0.11	18.87	1.48	1.35	0.08
	Pri comp/sec inc	50.47	3.28	0.87	0.07	46.34	1.79	0.99	0.04
	Sec comp and higher	23.16	3.05	0.95	0.13	34.79	1.74	1.02	0.05
1994-98	No edu/pri inc					21.05	3.76	1.37	0.18
	Pri comp/sec inc					18.55	3.45	1.07	0.19
	Sec comp and higher					60.41	4.65	1.28	0.08

Table 4.4 Changes in standard errors for estimates of educational attainment

		1993			1998		2003
		1 survey added	2 surveys added	3 surveys added	1 survey added	2 surveys added	1 survey added
1949-53	No edu/pri inc	0.9	-	-			
	Pri comp/sec inc	0.3	-	-			
	Sec comp and higher	1.7	-	-			
1954-58	No edu/pri inc	24.0	-	-			
	Pri comp/sec inc	21.7	-	-			
	Sec comp and higher	20.6	-	-			
1959-63	No edu/pri inc	24.3	38.4	38.5	31.6	-	0.1
	Pri comp/sec inc	28.6	39.0	39.1	24.4	-	0.1
	Sec comp and higher	22.6	33.6	33.7	27.2	-	0.2
1964-68	No edu/pri inc	31.6	45.1	51.8	29.9	41.3	26.4
	Pri comp/sec inc	30.8	39.9	46.9	21.8	34.9	29.7
	Sec comp and higher	15.7	26.6	32.6	28.0	38.5	25.2
1969-73	No edu/pri inc	35.3	48.6	53.9	29.0	38.1	22.7
	Pri comp/sec inc	26.3	41.2	47.6	32.5	42.9	23.7
	Sec comp and higher	14.2	33.6	41.5	37.4	48.5	22.7
1974-78	No edu/pri inc	41.1	52.0	58.0	22.3	33.9	28.2
	Pri comp/sec inc	26.0	41.4	48.0	31.9	42.8	24.0
	Sec comp and higher	17.1	34.7	39.2	33.6	41.4	17.0
1979-83	No edu/pri inc	53.1	65.1	72.5	29.9	46.1	34.2
	Pri comp/sec inc	44.7	54.7	61.8	24.5	38.1	29.4
	Sec comp and higher	24.9	34.7	40.8	21.9	31.5	24.1
1984-88	No edu/pri inc				55.3	69.8	36.5
	Pri comp/sec inc				47.1	59.9	29.4
	Sec comp and higher				42.5	48.5	15.3
1989-93	No edu/pri inc						50.6
	Pri comp/sec inc						45.4
	Sec comp and higher						42.9
1994-98	No edu/pri inc						
	Pri comp/sec inc						
	Sec comp and higher						

Contraceptive use

The second variable tested was contraceptive use, which was used exactly as it was in the datasets with four categories: None, folkloric, traditional and modern. This variable, unlike the educational attainment variable, is one that is expected to change over time for each cohort, because contraceptive behavior changes with age. Therefore combined cohort estimates yield statistics that are not useful. However, proportions for the same age groups do not stay constant either, because there is an increasing trend of contraceptive use with increasing age. Combined averages for the same age groups over time show the average contraceptive use over the survey periods 1993-2008, 1998-2008 and 2003-2008 (Tables 4.5, 4.6 and 4.7).

Table 4.5 Estimates of contraceptive methods and their standard errors from cumulations of TDHS-1993, TDHS-1998, TDHS-2003 and TDHS-2008

	1993				1993&1998				1993,1998&2003				All 4			
	Percent	SE	DEFF	CV	Percent	SE	DEFF	CV	Percent	SE	DEFF	CV	Percent	SE	DEFF	CV
None	39.73	0.809	1.638	0.020	39.28	0.559	1.403	0.014	36.79	0.420	1.276	0.011	35.07	0.363	1.244	0.010
Folkloric	0.81	0.115	0.963	0.143	0.71	0.082	0.986	0.116	0.59	0.061	1.027	0.103	0.46	0.047	1.017	0.102
Traditional	26.18	0.730	1.699	0.028	25.20	0.526	1.566	0.021	25.74	0.402	1.430	0.016	25.67	0.351	1.377	0.014
Modern	33.28	0.710	1.377	0.021	34.81	0.542	1.375	0.016	36.88	0.409	1.191	0.011	38.80	0.355	1.113	0.009

Table 4.6 Estimates of contraceptive methods and their standard errors from cumulations of TDHS-1998, TDHS-2003 and TDHS-2008

	1998				1998&2003				1998, 2003 & 2008			
	Proportion	SE	DEFF	CV	Proportion	SE	DEFF	CV	Proportion	SE	DEFF	CV
None	38.89	0.773	1.254	0.020	35.54	0.489	1.148	0.014	33.80	0.404	1.159	0.012
Folkloric	0.62	0.116	1.000	0.188	0.49	0.071	1.053	0.144	0.37	0.051	1.036	0.138
Traditional	24.35	0.753	1.490	0.031	25.56	0.482	1.346	0.019	25.53	0.400	1.318	0.016
Modern	36.14	0.804	1.379	0.022	38.40	0.499	1.145	0.013	40.30	0.409	1.078	0.010

Table 4.7 Estimates of contraceptive methods and their standard errors from cumulations of TDHS-2003 and TDHS-2008

	2003				2003 & 2008			
	Proportion	SE	DEFF	CV	Proportion	SE	DEFF	CV
None	32.43	0.601	1.001	0.019	31.48	0.467	1.099	0.015
Folkloric	0.38	0.084	1.147	0.223	0.26	0.052	1.077	0.203
Traditional	26.69	0.610	1.179	0.023	26.06	0.470	1.236	0.018
Modern	40.50	0.607	0.915	0.015	42.20	0.469	0.956	0.011

The point estimates presented in Tables 4.5, 4.6 and 4.7 show the cross-sectional proportions of type of contraceptive use at different survey dates. According to Table 4.5, the proportion of women who are not using any method at survey date decreased, proportions of folkloric and traditional methods decreased only slightly, and the proportion of modern method users increased from TDHS-1993 to TDHS-2008. The combined estimates from TDHS-1993 and TDHS-1998 provided a value between those obtained from these two surveys separately – as expected – and the standard errors are lower. The coefficient of variation is halved for none users and traditional users from 1993 to all four surveys combined. Design effects also decreased.

The combined estimates from TDHS-1993 and TDHS-1998 are the average proportions corresponding to the period 1993-1998. The latter combined estimates of three and four surveys have even lower standard errors, yet they correspond to longer periods of time; 10 and 15 years respectively. The highest decreases in standard errors were observed with the addition of one dataset only (column 1 in Table 4.8). The gain in standard errors decreased as more datasets were added. The coefficient of variation and design effect decreased when more surveys were added to TDHS-1998.

The gains in standard errors by adding one more survey is greater for TDHS-1998 (with the addition of 2003 data) compared to TDHS-1993 (with the addition of 1998 data). The standard errors are almost halved when both TDHS-2003 and

TDHS-2008 are added to TDHS-1998 (Table 4.8). For TDHS-2003, the addition of TDHS-2008 data has decreased the standard error by about 23 percent for all contraceptive methods, except for folkloric methods. The design effect increased in some cases (Table 4.7).

Table 4.8 Changes in standard errors for estimates of contraceptive use

	1993			1998		2003
	1 survey added	2 surveys added	3 surveys added	1 survey added	2 surveys added	1 survey added
None	30.9	48.0	55.2	36.8	47.7	22.3
Folkloric	28.9	47.5	59.2	38.8	56.0	38.2
Traditional	27.9	44.9	51.9	35.9	46.8	22.9
Modern	23.7	42.4	50.0	37.9	49.2	22.8

4.1.2. Means: Ideal Mean Number of Children and Mean Number of Children Ever Born

Both examples for means are presented in terms of birth cohorts. While the complex sample descriptive first mean (ideal number of children) can be obtained directly from the built in procedures of SAS, a SAS macro had to be written for the other mean (total children ever born) because the numerator and denominator are defined from different populations, as explained in the methodology section.

Ideal Number of Children Ever Born

The ideal number of children represents hypothetical values stated by women at survey date. As mentioned in Section 2.4.3, this number may change for cohorts at each survey. Thus at any survey date, it would show the ideal for a given cohort at that point in time only. The combined estimates show the average ideal number of children for cohorts in corresponding age groups, giving equal weight to cohorts at each survey; so that they contribute equally to the overall estimates as explained in the methodology section. This section only presents results when all surveys in the Tables provide data on birth cohorts.

Table 4.9 shows the results from adding datasets to TDHS-1993. The point estimates increase with the addition of each survey, suggesting an increase in the ideal number of children for cohorts. For the 1964-1968 and 1969-73 birth cohorts, this increase is significant with a non-overlapping confidence intervals from TDHS-1993 to TDHS-1993 and TDHS-1998 combined. The mean number of ideal number of children is more stable for the younger two cohorts. The gain in standard errors is the highest for these younger cohorts (Table 4.12). It is reduced by 60% for the 1979-1983 cohort after the addition of data from 3 more surveys to TDHS-1993, where the only apparent decrease in the coefficient of variation was observed (Table 4.9).

The ideal number of children estimates obtained from adding later datasets to TDHS-1998 (Table 4.10) resulted in more consistent means compared to those in Table 4.9. The confidence intervals of the means in TDHS-1998 are overlapping with the later means. The coefficient of variation decreased for all birth cohorts except for the 1979-1983 birth cohort. The design effects also decreased for these birth cohorts. The gains in standard errors range from 30 percent to 44 percent after the addition of TDHS-2003 and TDHS-2008 to TDHS-1998.

The addition of TDHS-2008 data to TDHS-2003 resulted in estimates that are similar to those obtained from TDHS-2003 only, however, the previously observed trend of slightly increasing point estimates remained (Table 4.11). The confidence intervals of respective estimates for birth cohorts overlap, suggesting that these increases are not significant. The gain in standard errors was the highest for the youngest two cohorts (Table 4.12). The coefficient of variation decreased or remained similar for all birth cohorts (Table 4.11). The design effect decreased for the first three (older) cohorts, and decreased for the younger three cohorts.

Table 4.9 Ideal number of children for 5 year birth cohorts and its standard error from cumulations of TDHS-1993, TDHS-1998, TDHS-2003 and TDHS-2008

	1993				1993 and 1998				1993, 1998 and 2003				All four surveys			
	R	SE	DEFF	CV	R	SE	DEFF	CV	R	SE	DEFF	CV	R	SE	DEFF	CV
1964-68	2.44	0.031	0.973	0.01	2.56	0.028	1.049	0.01	2.57	0.024	1.029	0.01	2.63	0.023	1.025	0.01
1969-73	2.33	0.027	1.008	0.01	2.45	0.027	1.271	0.01	2.48	0.024	1.229	0.01	2.54	0.021	1.130	0.01
1974-78	2.28	0.030	1.033	0.01	2.34	0.024	1.206	0.01	2.38	0.021	1.153	0.01	2.44	0.019	1.077	0.01
1979-83	2.30	0.055	1.041	0.02	2.35	0.031	1.008	0.01	2.35	0.026	1.007	0.01	2.40	0.022	1.050	0.01

Table 4.10 Ideal number of children for 5 year birth cohorts and its standard error from cumulations of TDHS-1998, TDHS-2003 and TDHS-2008

	1998				1998 and 2003				1998, 2003 and 2008			
	R	SE	DEFF	CV	R	SE	DEFF	CV	R	SE	DEFF	CV
1964-68	2.69	0.05	1.105	0.02	2.66	0.03	1.054	0.01	2.70	0.03	1.027	0.01
1969-73	2.56	0.05	1.413	0.02	2.58	0.03	1.323	0.01	2.63	0.03	1.166	0.01
1974-78	2.39	0.04	1.335	0.02	2.44	0.03	1.215	0.01	2.51	0.02	1.075	0.01
1979-83	2.41	0.03	0.932	0.01	2.39	0.02	0.945	0.01	2.45	0.02	1.066	0.01
1984-88	2.34	0.04	1.318	0.02	2.39	0.03	1.249	0.01	2.40	0.02	1.222	0.01

Table 4.11 Ideal number of children for 5 year birth cohorts and its standard error from cumulations of TDHS-2003 and TDHS-2008

	2003				2003 and 2008			
	R	SE	DEFF	CV	R	SE	DEFF	CV
1964-68	2.61	0.05	0.939	0.02	2.70	0.04	0.982	0.01
1969-73	2.62	0.04	1.000	0.02	2.67	0.03	0.983	0.01
1974-78	2.54	0.04	0.983	0.02	2.58	0.03	0.930	0.01
1979-83	2.34	0.03	0.987	0.01	2.47	0.03	1.142	0.01
1984-88	2.49	0.05	1.109	0.02	2.44	0.03	1.152	0.01
1989-93	2.39	0.06	1.047	0.03	2.41	0.04	1.172	0.02

Table 4.12 Changes in standard errors for estimates of ideal number of children

	1993			1998		2003
	1 survey added	2 surveys added	3 surveys added	1 survey added	2 surveys added	1 survey added
1964-68	9.0	20.7	24.6	25.7	36.6	21.1
1969-73	-1.2	13.2	24.1	27.2	44.2	27.4
1974-78	18.7	28.5	36.3	23.0	38.2	33.0
1979-83	42.6	52.3	59.7	23.4	30.1	14.7
1984-88				22.5	40.6	38.9
1989-93						40.3

Mean Number of Children Ever Born

The mean number of children ever born is presented for all women aged 15-49 and for women aged 40-49 in the most recent Turkey DHS report. This indicator is calculated for five year birth cohorts and as a weighted mean of these birth cohorts for all women in reproductive ages in this section. The results in the main reports are presented for the survey date, however, they are based on selected dates for this section; and are calculated from the birth history section rather than the single total number of children variable. The reason is to allow the addition of more surveys with a standard reference date.

For TDHS-1993, the mean was calculated from the number of children born up to September 1993. Results show that (Table 4.13) additional dataset reduced the standard error of the estimates for cohorts. The point estimates mostly increase as more surveys are added for almost all birth cohorts. This increase is not significant for most cases, however, the existence of an increasing trend suggest a small degree of bias. For the youngest cohort in TDHS-1993 for example, the mean number of children ever born is 0.06 in according to this survey; and is 0.11 for all four surveys combined. This is a significant increase according to the standard errors, and may mean that children or women with children were underreported in the earliest survey. It may also point out to recall errors from the newer surveys. The design effects tend to decrease with each additional survey to TDHS-1993, and this reverses with the addition of TDHS-2008; after which the design effects increase again. The coefficients of variation decrease in line with decreasing standard errors.

The overall mean number of children ever born is calculated according to the age structure of the TDHS-1993 survey. The point estimate increased by 0.03 after combining all datasets, with the standard error decreasing by 0.003 only. The decrease in the coefficient of variation is exaggerated due to rounding.

The results as of September, 1998 are provided from TDHS-1998, TDHS-2003 and TDHS-2008 (Table 4.14). Point estimates are more stable compared to those in Table 4.13. Only for the 1969-1973 birth cohort, the mean number of children increases by almost 0.1; from 1.66 to 1.75; which is not significant in terms of the confidence intervals. The design effects showed rather irregular changes with additional changes.

Table 4.13 The mean number of children ever born to birth cohorts of women as of September, 1993 and its standard error

	1993				1993 and 1998				1993, 1998 and 2003				1993, 1998, 2003 and 2008			
	R	SE	DEFF	CV	R	SE	DEFF	CV	R	SE	DEFF	CV	R	SE	DEFF	CV
1949-53	4.41	0.11	1.23	0.02	4.45	0.08	1.12	0.02								
1954-58	3.80	0.08	1.20	0.02	3.80	0.06	1.15	0.02	3.71	0.05	1.07	0.01				
1959-63	2.93	0.06	1.35	0.02	2.93	0.04	1.22	0.02	2.94	0.04	1.11	0.01	2.94	0.03	1.23	0.01
1964-68	1.84	0.04	1.42	0.02	1.89	0.04	1.53	0.02	1.91	0.03	1.18	0.02	1.92	0.03	1.23	0.01
1969-73	0.74	0.02	1.05	0.03	0.78	0.02	1.07	0.03	0.80	0.02	1.07	0.02	0.82	0.02	1.15	0.02
1974-78	0.06	0.01	1.55	0.17	0.09	0.01	1.31	0.10	0.10	0.01	1.13	0.07	0.11	0.01	1.20	0.06
Overall	2.04	0.07	1.41	0.04	2.06	0.07	1.42	0.04	2.06	0.07	1.19	0.03	2.07	0.07	1.42	0.03

Table 4.14 The mean number of children ever born to birth cohorts of women as of September, 1998 and its standard error

	1998				1998 and 2003				1998, 2003 and 2008			
	R	SE	DEFF	CV	R	SE	DEFF	CV	R	SE	DEFF	CV
1954-58	3.92	0.093	1.163	0.02	3.78	0.062	1.165	0.02				
1959-63	3.25	0.072	1.128	0.02	3.28	0.050	1.241	0.02	3.28	0.042	1.268	0.01
1964-68	2.54	0.065	1.548	0.03	2.58	0.042	1.406	0.02	2.58	0.034	1.234	0.01
1969-73	1.66	0.047	1.379	0.03	1.72	0.034	1.450	0.02	1.75	0.027	1.371	0.02
1974-78	0.72	0.028	1.079	0.04	0.74	0.019	1.083	0.03	0.74	0.017	1.086	0.02
1979-83	0.08	0.009	1.100	0.10	0.08	0.006	1.225	0.07	0.08	0.005	1.122	0.06
Overall	2.00	0.027	1.138	0.01	2.01	0.022	0.973	0.01	2.02	0.021	0.917	0.01

The cohort values obtained from TDHS-1998 and later surveys as of September 1998 (Table 4.14) are generally higher compared to those found in the previous table (Table 4.13), given the five year difference in between. Other than this overall increase, the results for September 1998 are similar to those of September 1993. With each additional dataset, the point estimates change by a small amount, however, there is substantial decrease in the standard errors. The overall mean number of children estimate has a higher decrease in standard errors for September 1998 compared to September 1993.

Table 4.15 shows the single and combined data estimates with a time limit of December, 2003. The point estimates are generally quite consistent, and the standard errors are on average 25 percent lower for the combined data set estimates (excluding the 1979-1984 birth cohort, which is 12 percent lower).

Table 4.15 Mean number of children for 5 year birth cohorts and its standard error from cumulations of TDHS-2003 and TDHS-2008

	2003				2003 and 2008			
	R	SE	DEFF	CV	R	SE	DEFF	CV
1959-63	3.43	0.070	1.365	0.02	3.42	0.054	1.329	0.02
1964-68	2.95	0.057	1.188	0.02	2.96	0.043	1.155	0.01
1969-73	2.40	0.053	1.448	0.02	2.42	0.037	1.353	0.02
1974-78	1.56	0.035	1.291	0.02	1.58	0.029	1.298	0.02
1979-84	0.60	0.019	1.156	0.03	0.63	0.017	1.305	0.03
1989-94	0.07	0.009	0.680	0.14	0.08	0.007	0.788	0.09
Overall	1.83	0.049	0.975	0.03	1.85	0.050	1.089	0.03

Table 4.16 Changes in standard errors for estimates of mean number of children ever born

	1993			1998		2003
	1 survey added	2 surveys added	3 surveys added	1 survey added	2 surveys added	1 survey added
1949-53	29.3					
1954-58	26.5	41.7		34.0		
1959-63	25.1	39.5	45.5	31.0	41.3	23.5
1964-68	13.3	32.4	40.8	35.6	48.3	24.8
1969-73	15.0	27.3	34.2	28.5	43.0	30.2
1974-78	15.0	26.8	34.2	31.8	40.9	19.0
1979-83				31.3	42.6	11.5
1984-88						27.4
Overall	1.4	2.8	3.4	18.7	23.0	-3.2

4.2. Combined Data Estimates of Rates

4.2.1 Total Fertility Rate

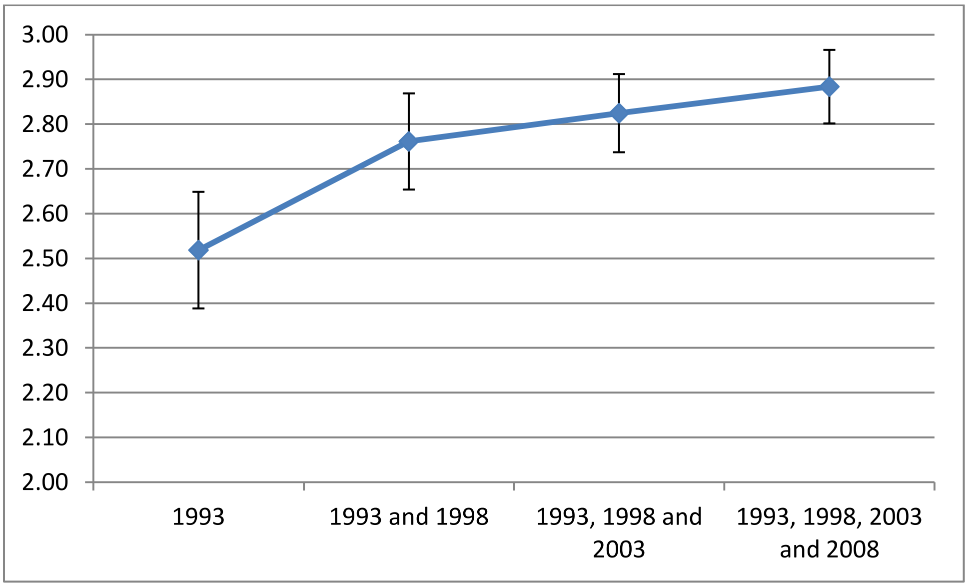
The TFR is usually calculated for a three year period of time in Turkish DHSs. The exception is TDHS-1998, where the main report presented a one year TFR instead. This section re-calculates these rates with the addition of extra data sets. The addition of extra data sets was done for the overall sample, as well as some basic variables used in TDHS reports for the presentation of this indicator: Type of place of residence, region and education. Two and one year TFRs were also calculated to see how standard errors change with shorter durations. These are presented in less detail (estimate and standard error), and for one variable only (region).

TDHS-1993

In order to compute TFRs corresponding to a three year period prior to TDHS-1993, an artificial date was selected to be used as interview date, as was for the mean number of children in the previous section. The earliest interview month of TDHS-1993 was selected for this purpose, which was September, so that no cases from this survey would be left out. The birth history sections of all remaining three surveys were used up to this date only.

The TFR calculated for Turkey goes up by 0.24 with the addition of TDHS-1998 to TDHS-1993 (Table 4.17). The later surveys also cause an increase, but not as dramatic. The standard errors, on the other hand, decrease as expected. The overall decrease observed in the standard error is 37 percent. While the confidence intervals for the last three estimates overlap, the one for the TFR computed from TDHS-1998 does not, implying this estimate is significantly lower than the other three (Figure 4.1).

Figure 4.1 Three year TFR for Turkey prior to September 1993 and confidence intervals



For urban areas, an 11% increase in TFR observed from TDHS-1993 to all surveys combined. There is a difference of about 0.3 children per woman between TDHS-1993 and all surveys combined. The overall decrease in standard error is 41 percent. The increase in point estimates is higher for rural areas, the TFR increases by 19 percent and the decrease in standard error is lower. Both for urban and rural areas; the confidence intervals for the first and last estimates are non-overlapping.

Among regions; the highest change in TFR as a result of combining datasets is the region with the highest TFR to begin with; the Eastern region. The TFR for this region increases by 28 percent, reaching 5.22 children per woman when all surveys are combined. The largest gain in standard errors is observed for the Central region. The most consistent estimate was obtained for the Northern region. For the Northern and Central regions, all four confidence intervals overlap. For the remaining three regions, results suggest significantly differed for the first and last estimates.

The lowest and middle education categories showed increases of 15 percent and 19 percent respectively for the TFR estimate comparisons of last to first. For the highest education category, all four confidence intervals overlap, despite an increase in TFR with additional datasets. The combined estimates suggest 4.6 and 1.7 children per woman for the lowest and highest education levels, respectively.

Table 4.17 Three-year total fertility rates prior to September, 1993, Turkey, type of place of residence, regions and educational attainment

Turkey		TFR	SE(TFR)	DEFF	CV	Decrease in SE (percent)
	1993	2.52	0.065	1.538	0.03	
	1993 and 1998	2.76	0.054	1.462	0.02	17
	1993, 1998 and 2003	2.82	0.044	1.343	0.02	32
	1993, 1998, 2003 and 2008	2.88	0.041	1.329	0.01	37

Type of place of residence		TFR	SE(TFR)	DEFF	CV	Decrease in SE (percent)
Urban						
	1993	2.31	0.076	1.608	0.03	
	1993 and 1998	2.52	0.059	1.362	0.02	22
	1993, 1998 and 2003	2.56	0.047	1.241	0.02	38
	1993, 1998, 2003 and 2008	2.63	0.045	1.248	0.02	41
Rural						
	1993	2.89	0.126	1.545	0.04	
	1993 and 1998	3.23	0.119	1.852	0.04	6
	1993, 1998 and 2003	3.37	0.098	1.659	0.03	22
	1993, 1998, 2003 and 2008	3.43	0.090	1.593	0.03	29

Table 4.17 (continued). Three-year total fertility rates prior to September, 1993, regions

Region		TFR	SE(TFR)	DEFF	CV	Decrease in SE (percent)
West						
	1993	1.92	0.075	0.960	0.04	
	1993 and 1998	2.08	0.078	1.284	0.04	-4
	1993, 1998 and 2003	2.19	0.06	1.057	0.03	20
	1993, 1998, 2003 and 2008	2.29	0.062	1.219	0.03	17
South						
	1993	2.35	0.141	1.885	0.06	
	1993 and 1998	2.65	0.106	1.474	0.04	25
	1993, 1998 and 2003	2.81	0.093	1.418	0.03	34
	1993, 1998, 2003 and 2008	2.85	0.082	1.272	0.03	42
Central						
	1993	2.45	0.144	1.766	0.06	
	1993 and 1998	2.56	0.102	1.385	0.04	29
	1993, 1998 and 2003	2.49	0.084	1.311	0.03	42
	1993, 1998, 2003 and 2008	2.55	0.075	1.250	0.03	48
North						
	1993	2.74	0.137	1.080	0.05	
	1993 and 1998	2.75	0.098	0.799	0.04	28
	1993, 1998 and 2003	2.74	0.095	1.032	0.03	31
	1993, 1998, 2003 and 2008	2.76	0.091	1.059	0.03	34
East						
	1993	4.08	0.270	1.863	0.07	
	1993 and 1998	4.92	0.223	1.812	0.05	17
	1993, 1998 and 2003	5.09	0.185	1.682	0.04	31
	1993, 1998, 2003 and 2008	5.22	0.173	1.600	0.03	36

Table 4.17 (continued). Three-year total fertility rates prior to September, 1993, educational attainment

Education	TFR	SE(TFR)	DEFF	CV	Decrease in SE (percent)
No education or first level primary shool incomplete					
1993	3.85	0.143	1.184	0.04	
1993 and 1998	4.33	0.12	1.243	0.03	16
1993, 1998 and 2003	4.48	0.098	1.173	0.02	31
1993, 1998, 2003 and 2008	4.59	0.086	1.030	0.02	40
First level primary complete, second level incomplete					
1993	2.31	0.067	1.082	0.03	
1993 and 1998	2.52	0.054	0.912	0.02	19
1993, 1998 and 2003	2.6	0.047	0.956	0.02	30
1993, 1998, 2003 and 2008	2.65	0.045	0.994	0.02	33
Second level+					
1993	1.56	0.084	0.962	0.05	
1993 and 1998	1.67	0.077	1.199	0.05	8
1993, 1998 and 2003	1.63	0.061	1.088	0.04	27
1993, 1998, 2003 and 2008	1.66	0.059	1.147	0.04	30

The same methodology was used to calculate two year TFRs from the combined data set. The same date – September 1993 - was used as the starting reference date. 2 years of birth history prior to this date was used to calculate TFR for each region and Turkey as a whole. This was repeated for 5 different periods back in time.

Figure 4.2 Two year TFRs preceding September 1993

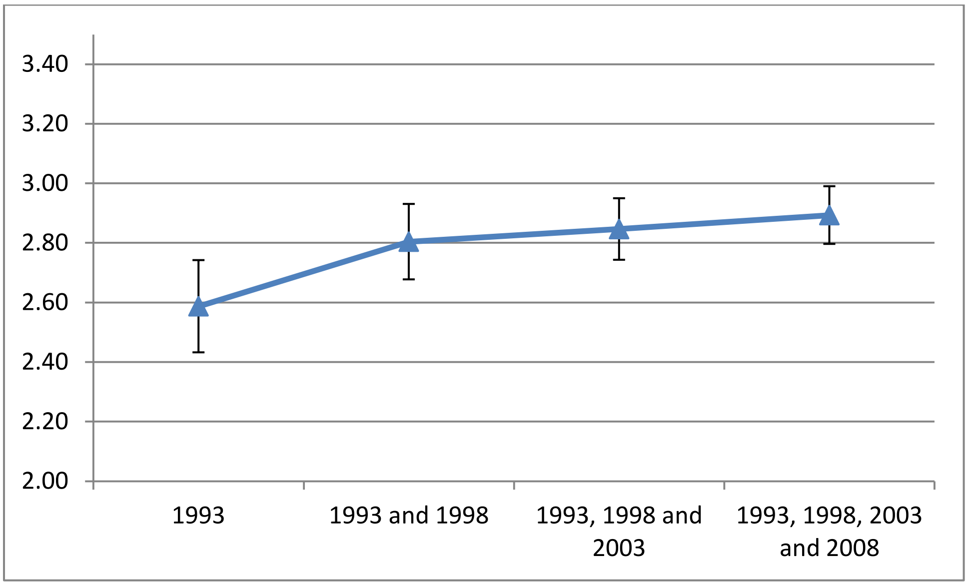


Table 4.18 2 year TFRs prior to September 1993; 1991 and 1989

		September 1993				September 1991				September 1989			
		TFR	SE(TFR)	DEFF	CV	TFR	SE(TFR)	DEFF	CV	TFR	SE(TFR)	DEFF	CV
Turkey	1993	2.59	0.077	1.445	0.03	2.54	0.070	1.161	0.03	3.12	0.252	0.952	0.08
	1993 and 1998	2.80	0.063	1.419	0.02	2.80	0.055	1.066	0.02	3.21	0.245	0.935	0.08
	1993, 1998 and 2003	2.85	0.052	1.317	0.02	2.91	0.047	1.051	0.02	3.22	0.243	0.929	0.08
	1993, 1998, 2003 and 2008	2.89	0.048	1.314	0.02	2.96	0.044	1.027	0.01	3.27	0.242	0.927	0.07
West	1993	1.95	0.094	0.955	0.05	1.93	0.100	0.955	0.05	2.75	0.574	0.800	0.21
	1993 and 1998	2.17	0.100	1.359	0.05	2.05	0.082	0.924	0.04	2.88	0.574	0.807	0.20
	1993, 1998 and 2003	2.28	0.077	1.132	0.03	2.17	0.069	0.925	0.03	2.84	0.572	0.806	0.20
	1993, 1998, 2003 and 2008	2.33	0.075	1.223	0.03	2.23	0.065	0.934	0.03	2.91	0.572	0.805	0.20
South	1993	2.44	0.177	1.968	0.07	2.49	0.133	0.974	0.05	2.96	0.183	1.308	0.06
	1993 and 1998	2.59	0.128	1.509	0.05	2.87	0.105	0.874	0.04	3.18	0.140	1.153	0.04
	1993, 1998 and 2003	2.76	0.113	1.442	0.04	2.99	0.093	0.863	0.03	3.22	0.124	1.102	0.04
	1993, 1998, 2003 and 2008	2.79	0.099	1.288	0.04	3.07	0.088	0.885	0.03	3.24	0.116	1.065	0.04
Central	1993	2.45	0.159	1.501	0.07	2.53	0.147	1.311	0.06	2.76	0.194	1.771	0.07
	1993 and 1998	2.53	0.112	1.145	0.04	2.66	0.121	1.260	0.05	2.71	0.133	1.401	0.05
	1993, 1998 and 2003	2.42	0.095	1.199	0.04	2.76	0.102	1.154	0.04	2.75	0.109	1.234	0.04
	1993, 1998, 2003 and 2008	2.51	0.087	1.177	0.03	2.78	0.096	1.169	0.03	2.83	0.101	1.195	0.04
North	1993	2.80	0.170	1.176	0.06	2.71	0.156	0.898	0.06	2.77	0.215	1.339	0.08
	1993 and 1998	2.77	0.124	0.922	0.04	2.89	0.132	0.912	0.05	2.78	0.160	1.346	0.06
	1993, 1998 and 2003	2.68	0.111	1.076	0.04	2.94	0.120	0.965	0.04	2.90	0.137	1.225	0.05
	1993, 1998, 2003 and 2008	2.68	0.104	1.093	0.04	2.97	0.111	0.941	0.04	2.89	0.125	1.142	0.04
East	1993	4.30	0.330	1.831	0.08	4.01	0.242	1.123	0.06	4.80	0.318	1.275	0.07
	1993 and 1998	5.05	0.261	1.817	0.05	4.86	0.189	1.029	0.04	5.13	0.233	1.192	0.05
	1993, 1998 and 2003	5.17	0.215	1.683	0.04	5.11	0.163	1.064	0.03	5.24	0.205	1.165	0.04
	1993, 1998, 2003 and 2008	5.29	0.201	1.614	0.04	5.22	0.153	1.028	0.03	5.30	0.196	1.138	0.04

The two year estimates are fairly close to the 3 year estimates. The three year TFR for Turkey from the TDHS-1993 data set was calculated as 2.52 from September 1993 and backwards (Table 4.17), and the corresponding two year TFR was 2.59 (Table 4.18): Slightly higher, but not significantly different. This still holds after the addition of more data sets. Among regions, the two year TFRs are slightly higher for two of the regions; East and West. The TFR mostly decreases when calculated from two years instead of three, in the remaining regions.

The observation of higher TFRs with the addition of more data sets is still valid (Figure 4.2) based on the results from Turkey as a whole. The confidence intervals of the estimate obtained from TDHS-1993 by itself and from all four surveys combined do not overlap.

As expected, the standard errors are higher for the 2 year TFRs than the 3 year TFRs. As far as the gains, they were similar for 3 and 2 year TFRs. The standard error of the 2 year TFR preceding September 1993 was decreased by 28 percent after the addition of three more data sets.

Generally, TFRs increased as calculations were performed for older dates. This agrees with the fact that fertility has been declining in Turkey. However, an unexpected fluctuation is observed for the West region. From 1993 to 1991, it decreases; to 1989 it increases again. Yet again, these differences are not significant, the confidence intervals overlap. The sizes of standard errors are striking for the West region in 1989, compared to 1991, affecting the overall standard error as well. They are more than 5 times as big in 1989. The estimates for TFR in 1989 yield very large confidence intervals and these do not shrink with additional data sets.

The single year TFRs are generally lower than the two year TFRs after combining all surveys (except for the South and North regions, Table 4.19). For Turkey, the standard errors for the single year TFR preceding September 1992 is higher than both that of September 1993 and September 1991. Looking at regions, the East is the most responsible for this. Point estimates increase in almost all cases, and standard errors decrease with additional surveys with a few exceptions.

Table 4.19 Single year TFRs prior to September 1993; 1992 and 1991

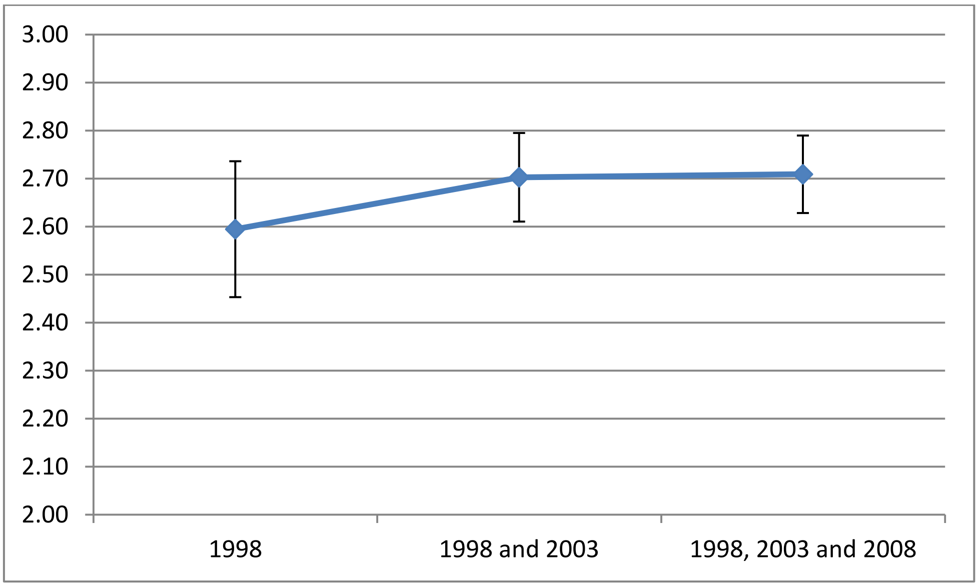
		September 1993				September 1992				September 1991			
		TFR	SE(TFR)	DEFF	CV	TFR	SE(TFR)	DEFF	CV	TFR	SE(TFR)	DEFF	CV
Turkey	1993	2.69	0.101	1.175	0.04	2.48	0.115	1.441	0.09	2.37	0.089	0.983	0.04
	1993 and 1998	2.73	0.077	1.102	0.03	2.89	0.186	1.373	0.06	2.67	0.076	1.032	0.03
	1993, 1998 and 2003	2.79	0.065	1.102	0.02	2.91	0.150	1.261	0.05	2.78	0.064	1.009	0.02
	1993, 1998, 2003 and 2008	2.79	0.065	1.102	0.02	2.91	0.150	1.261	0.05	2.86	0.061	1.015	0.02
West	1993	2.05	0.139	0.945	0.07	1.84	0.132	0.840	0.07	1.86	0.137	0.901	0.07
	1993 and 1998	2.20	0.130	1.107	0.06	2.15	0.137	1.183	0.06	1.88	0.114	0.882	0.06
	1993, 1998 and 2003	2.28	0.108	1.074	0.05	2.28	0.109	1.056	0.05	2.00	0.093	0.858	0.05
	1993, 1998, 2003 and 2008	2.27	0.102	1.119	0.05	2.40	0.102	1.050	0.04	2.18	0.098	0.998	0.04
South	1993	2.54	0.232	1.428	0.09	2.34	0.188	1.062	0.08	2.16	0.189	1.097	0.09
	1993 and 1998	2.52	0.159	1.099	0.06	2.65	0.161	1.131	0.06	2.78	0.153	0.934	0.06
	1993, 1998 and 2003	2.82	0.146	1.105	0.05	2.70	0.144	1.188	0.05	2.90	0.140	1.006	0.05
	1993, 1998, 2003 and 2008	2.84	0.128	0.996	0.05	2.74	0.134	1.186	0.05	2.95	0.125	0.923	0.04
Central	1993	2.47	0.214	1.344	0.09	2.42	0.243	1.459	0.10	2.47	0.202	1.165	0.08
	1993 and 1998	2.42	0.151	1.116	0.06	2.65	0.170	1.146	0.06	2.63	0.167	1.216	0.06
	1993, 1998 and 2003	2.35	0.128	1.160	0.05	2.50	0.138	1.121	0.06	2.62	0.138	1.126	0.05
	1993, 1998, 2003 and 2008	2.43	0.115	1.084	0.05	2.59	0.128	1.087	0.05	2.64	0.126	1.100	0.05
North	1993	3.02	0.240	1.032	0.08	2.57	0.221	0.916	0.09	2.60	0.219	0.822	0.08
	1993 and 1998	2.95	0.188	0.958	0.06	2.58	0.172	0.923	0.07	2.71	0.194	0.944	0.07
	1993, 1998 and 2003	2.65	0.160	1.077	0.06	2.71	0.160	1.034	0.06	2.87	0.178	1.000	0.06
	1993, 1998, 2003 and 2008	2.75	0.146	1.034	0.05	2.59	0.144	1.050	0.06	2.91	0.167	0.979	0.06
East	1993	4.45	0.341	1.024	0.08	4.16	0.528	2.066	0.13	3.60	0.285	0.790	0.08
	1993 and 1998	4.55	0.241	0.99	0.05	5.58	0.423	1.868	0.08	4.63	0.266	1.022	0.06
	1993, 1998 and 2003	4.79	0.198	0.97	0.04	5.59	0.346	1.680	0.06	4.94	0.229	1.032	0.05
	1993, 1998, 2003 and 2008	4.89	0.182	0.94	0.04	5.74	0.323	1.587	0.06	5.07	0.214	0.995	0.04

TDHS-1998

The earliest interview month of TDHS-1998 was September, similar to TDHS-1993. Therefore 1998 September was set as the date prior to which the TFRs would be calculated. TDHS-1993 was excluded since it does not provide any data for the three years preceding September 1998. The birth history sections of all remaining surveys were used up to this date only.

The overall three year TFR for Turkey was estimated at 2.59 from TDHS-1998. The TFRs obtained by adding TDHS-2003, and TDHS-2003 and TDHS-2008 increased this estimate to 2.70 and 2.71 respectively. However, the confidence intervals for these estimates overlap (Figure 4.3), unlike the TFR calculated for TDHS-1993 in the previous section. This shows that overall estimate for 1998 is more stable than that of 1993.

Figure 4.3 Three year TFR for Turkey prior to September 1998 and confidence intervals



The overlapping confidence intervals phenomenon is observed for all subgroups in TDHS-1998. These subgroups are composed of women interviewed in different regions, places of residence and with various educational levels.

The TFR estimate for urban residences increases by 5 percent with the addition of two extra data sets, reaching 2.48. The rural TFR increase is 7 percent. The overall gains in standard errors are 39 percent and 42 percent respectively.

Similar to TDHS-1993, the highest change in the TFR estimate was obtained for the East region. The combined data estimate for North is very stable; which also was the case for the estimates of TDHS-1993. The TFR for the Central region is the only region portraying a decrease with additional datasets, although the decrease is not significant. The decrease in standard error is the highest for this region (43 percent).

The TFR for the lowest educational category of women showed the largest difference as a result of combining datasets. A very small difference was observed for the middle category, and a negative change was observed for the highest category. Similar to the Central region, the negative change is very small. The largest gain in standard errors was observed for the lowest educational level, followed by the highest (44 percent and 40 percent respectively).

Table 4.20 Three-year total fertility rates prior to September, 1998, Turkey, place of residence

Total		TFR	SE(TFR)	DEFF	CV	Decrease in SE (percent)
	1998	2.59	0.071	1.496	0.03	
	1998 and 2003	2.70	0.046	1.311	0.02	35
	1998, 2003 and 2008	2.71	0.040	1.329	0.01	44

Type of place of residence		TFR	SE(TFR)	DEFF	CV	Decrease in SE (percent)
Urban						
	1998	2.36	0.072	1.206	0.03	
	1998 and 2003	2.45	0.047	1.090	0.02	35
	1998, 2003 and 2008	2.48	0.044	1.232	0.02	39
Rural						
	1998	3.09	0.15	1.772	0.05	
	1998 and 2003	3.30	0.101	1.545	0.03	33
	1998, 2003 and 2008	3.30	0.084	1.462	0.03	44

Table 4.20 (continued). Three-year total fertility rates prior to September, 1998, Turkey, regions

Region		TFR	SE(TFR)	DEFF	CV	Decrease in SE (percent)
West						
	1998	2.00	0.114	1.399	0.06	
	1998 and 2003	2.15	0.07	1.119	0.03	39
	1998, 2003 and 2008	2.19	0.066	1.300	0.03	42
South						
	1998	2.59	0.148	1.313	0.06	
	1998 and 2003	2.60	0.112	1.520	0.04	24
	1998, 2003 and 2008	2.67	0.094	1.438	0.03	36
Central						
	1998	2.55	0.138	1.350	0.05	
	1998 and 2003	2.53	0.088	1.077	0.03	36
	1998, 2003 and 2008	2.50	0.078	1.153	0.03	43
North						
	1998	2.66	0.176	1.201	0.07	
	1998 and 2003	2.79	0.134	1.197	0.05	24
	1998, 2003 and 2008	2.66	0.113	1.195	0.04	36
East						
	1998	4.17	0.212	1.651	0.05	
	1998 and 2003	4.45	0.147	1.734	0.03	31
	1998, 2003 and 2008	4.52	0.121	1.573	0.03	43

Table 4.20 (continued). Three-year total fertility rates prior to September, 1998, Turkey, educational attianment

Education	TFR	SE(TFR)	DEFF	CV	Decrease in SE (percent)
No education or first level primary school incomplete					
1998	3.86	0.174	1.248	0.04	
1998 and 2003	4.27	0.123	1.311	0.03	29
1998, 2003 and 2008	4.36	0.098	1.156	0.02	44
First level primary complete, second level incomplete					
1998	2.59	0.09	1.378	0.03	
1998 and 2003	2.67	0.059	1.210	0.02	34
1998, 2003 and 2008	2.68	0.056	1.447	0.02	38
Second level and higher					
1998	1.80	0.112	1.219	0.06	
1998 and 2003	1.79	0.073	1.098	0.04	35
1998, 2003 and 2008	1.76	0.067	1.162	0.04	40

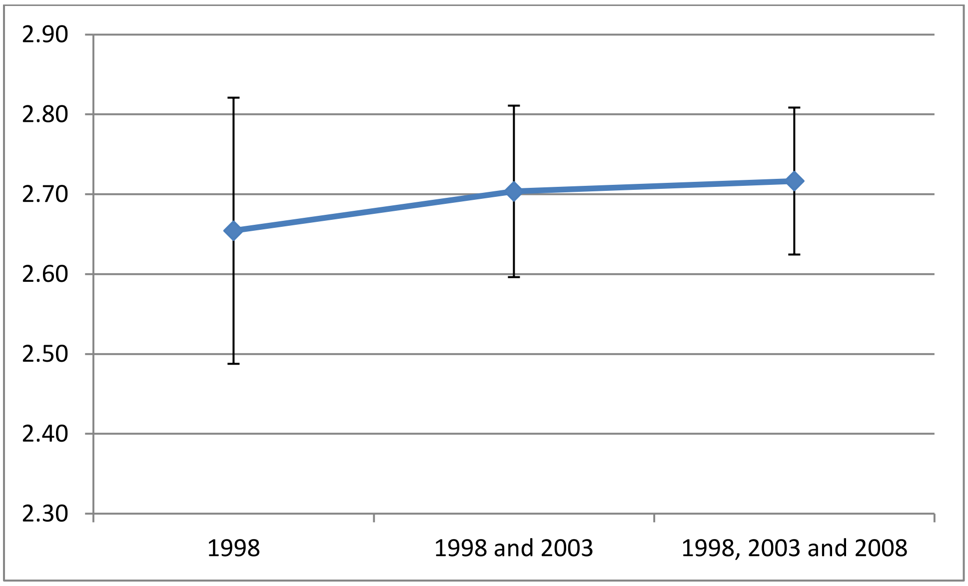
Table 4.21 Two-year total fertility rates prior to September; 1998, 1996 and 1994

		September 1998				September 1996				September 1994			
		TFR	SE(TFR)	DEFF	CV	TFR	SE(TFR)	DEFF	CV	TFR	SE(TFR)	DEFF	CV
Turkey	1998	2.65	0.085	1.408	0.03	2.53	0.089	1.548	0.04	2.74	0.105	0.922	0.04
	1998 and 2003	2.70	0.055	1.250	0.02	2.71	0.062	1.446	0.02	2.74	0.080	1.082	0.03
	1998, 2003 and 2008	2.72	0.047	1.210	0.02	2.76	0.053	1.360	0.02	2.73	0.069	1.031	0.03
West	1998	2.10	0.145	1.352	0.07	1.87	0.129	1.289	0.07	2.09	0.153	0.696	0.07
	1998 and 2003	2.11	0.085	1.062	0.04	2.09	0.088	1.175	0.04	2.23	0.125	0.928	0.06
	1998, 2003 and 2008	2.15	0.078	1.137	0.04	2.21	0.080	1.196	0.04	2.23	0.110	0.945	0.05
South	1998	2.56	0.183	1.486	0.07	2.59	0.171	1.088	0.07	2.76	0.233	1.022	0.08
	1998 and 2003	2.44	0.126	1.491	0.05	2.85	0.122	1.066	0.04	2.57	0.164	1.070	0.06
	1998, 2003 and 2008	2.59	0.107	1.425	0.04	2.87	0.104	1.056	0.04	2.70	0.138	0.966	0.05
Central	1998	2.62	0.170	1.318	0.06	2.41	0.179	1.522	0.07	2.82	0.213	0.813	0.08
	1998 and 2003	2.50	0.111	1.159	0.04	2.47	0.135	1.632	0.05	2.66	0.176	1.149	0.07
	1998, 2003 and 2008	2.52	0.092	1.087	0.04	2.50	0.112	1.528	0.04	2.51	0.141	1.012	0.06
North	1998	2.68	0.226	1.323	0.08	2.67	0.174	0.712	0.07	2.67	0.249	0.784	0.09
	1998 and 2003	2.78	0.171	1.350	0.06	2.48	0.121	0.773	0.05	2.38	0.176	0.882	0.07
	1998, 2003 and 2008	2.61	0.141	1.281	0.05	2.47	0.109	0.894	0.04	2.33	0.153	0.926	0.07
East	1998	4.20	0.245	1.617	0.06	4.39	0.343	2.136	0.08	4.44	0.342	1.119	0.08
	1998 and 2003	4.43	0.160	1.567	0.04	4.67	0.239	1.860	0.05	4.66	0.262	1.220	0.06
	1998, 2003 and 2008	4.49	0.131	1.429	0.03	4.77	0.206	1.635	0.04	4.77	0.237	1.197	0.05

Table 4.22 Single-year total fertility rates prior to September; 1998, 1997 and 1996

		September 1998				September 1997				September 1996			
		TFR	SE(TFR)	DEFF	CV	TFR	SE(TFR)	DEFF	CV	TFR	SE(TFR)	DEFF	CV
Turkey	1998	2.73	0.124	1.396	0.05	2.58	0.114	1.224	0.04	2.47	0.112	1.199	0.05
	1998 and 2003	2.72	0.080	1.268	0.03	2.68	0.079	1.214	0.03	2.70	0.077	1.108	0.03
	1998, 2003 and 2008	2.69	0.066	1.193	0.02	2.74	0.066	1.086	0.02	2.69	0.065	1.054	0.02
West	1998	2.15	0.199	1.152	0.09	2.03	0.215	1.470	0.11	1.80	0.160	0.917	0.09
	1998 and 2003	2.11	0.131	1.164	0.06	2.12	0.133	1.198	0.06	2.18	0.117	0.910	0.05
	1998, 2003 and 2008	2.11	0.108	1.104	0.05	2.20	0.111	1.053	0.05	2.22	0.101	0.910	0.05
South	1998	2.67	0.232	1.158	0.09	2.43	0.221	0.944	0.09	2.67	0.275	1.217	0.10
	1998 and 2003	2.48	0.165	1.272	0.07	2.40	0.153	0.995	0.06	2.75	0.187	1.168	0.07
	1998, 2003 and 2008	2.59	0.141	1.246	0.05	2.64	0.134	0.976	0.05	2.72	0.160	1.203	0.06
Central	1998	2.73	0.288	1.784	0.11	2.49	0.200	0.905	0.08	2.44	0.260	1.443	0.11
	1998 and 2003	2.61	0.175	1.319	0.07	2.38	0.157	1.085	0.07	2.48	0.180	1.279	0.07
	1998, 2003 and 2008	2.56	0.146	1.322	0.06	2.51	0.126	0.939	0.05	2.37	0.146	1.244	0.06
North	1998	2.76	0.336	1.371	0.12	2.62	0.267	0.876	0.10	2.61	0.241	0.706	0.09
	1998 and 2003	2.76	0.255	1.581	0.09	2.82	0.224	0.969	0.08	2.60	0.177	0.759	0.07
	1998, 2003 and 2008	2.60	0.206	1.521	0.08	2.67	0.192	0.967	0.07	2.60	0.149	0.776	0.06
East	1998	4.18	0.344	1.500	0.08	4.18	0.304	1.135	0.07	4.13	0.324	1.203	0.08
	1998 and 2003	4.32	0.215	1.368	0.05	4.55	0.227	1.440	0.05	4.41	0.225	1.260	0.05
	1998, 2003 and 2008	4.35	0.171	1.214	0.04	4.67	0.196	1.435	0.04	4.50	0.187	1.137	0.04

Figure 4.4 Two year TFRs for Turkey prior to September 1998 and confidence intervals



The combined data estimates for two year TFRs preceding September 1998 are very stable compared to the estimate obtained from TDHS-1998 only (Figure 4.4). This holds for regions as well. The next two year estimates prior to September 1996 are less stable, the TFR increased by 9 percent after two additional datasets. Only in the North region the TFR decreased with more data (decreasing from 2.67 to 2.47). The standard errors decreased with additional datasets without exceptions. The standard error of the TFR in the East region was higher than the other regions, which actually did not imply any less precision, looking at the coefficients of variation. The two year TFRs prior to September 1994 portrayed larger standard errors compared to later estimates. After combining surveys at this date, the largest coefficient of variation was 7 percent, still implying precise estimates.

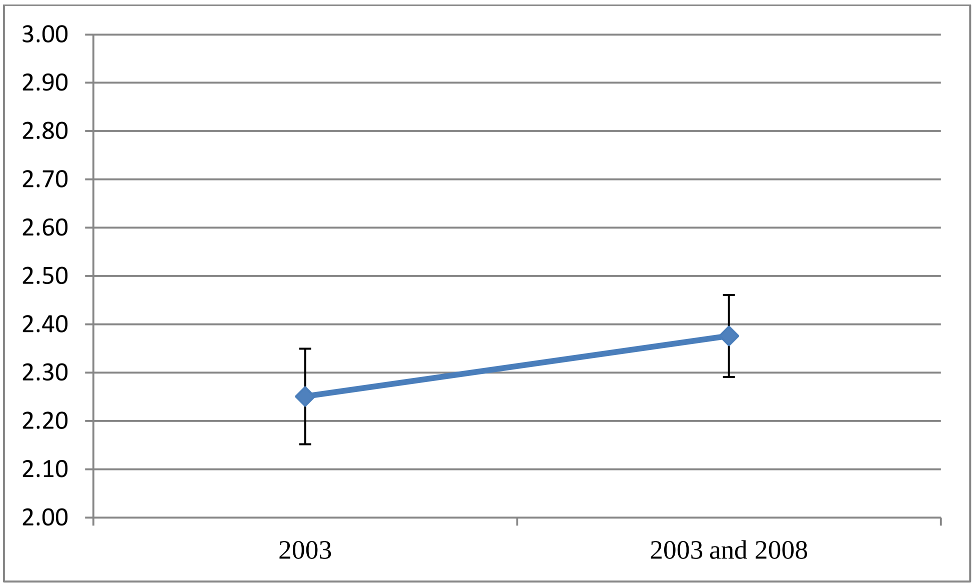
The single TFR year estimates preceding September 1998 are similar in level to the three year estimates (Table 4.22). They are stable for Turkey and the West region, and they decreased in the South, Central and North regions. These decreases were not significant. The coefficients of variation were generally larger for the one year estimates compared to the two and three year estimates. For the regions, it fell as low as 4 percent (in the East) from the combined dataset.

TDHS-2003

The fieldwork for TDHS-2003 started later in the year compared to TDHS-1993 and TDHS-1998. The earliest interview month of TDHS-2003 was December, and fieldwork ended in the following year. December 2003 was set as the date prior to which the TFRs would be calculated from this survey and TDHS-2008. Both TDHS-1993 and TDHS-1998 were excluded, since they did not have any data for the period of interest.

The overall three year TFR for Turkey was estimated at 2.25 from TDHS-2003. The TFR obtained by adding TDHS-2008 increased this estimate to 2.38, corresponding to a 6 percent increase. The standard error of the TFR estimate decreased by 14 percent after the extra data set, and the resulting confidence intervals overlapped with those of the original estimate.

Figure 4.5 Three year TFR for Turkey prior to September 2003 and confidence intervals



The increase in TFR as a result of the addition of TDHS-2008 was larger in rural areas than urban areas (Table 4.23), yet the gain in standard error was also higher in the former. None of the regional estimates significantly differed after the addition TDHS-2008, all estimates remained within overlapping confidence intervals. The least consistent estimate belonged to the West region with an increasing standard error, and the most stable to the South. A gain of 33 percent was obtained in the standard error for the South region, the highest among all regions.

The TFR estimates for different educational levels were also fairly stable after combining two surveys. The lowest education category had a TFR of 3.71 and the highest education category had a TFR of 1.50, these changed to 4.00 and 1.54 respectively, after combining TDHS-2008 with TDHS-2003. The reduction in the standard error was one third for the former, and about half of that for the latter (33 and 15 percent respectively).

TDHS-2003 also enables the estimation of total fertility rates in the basis of 12 regions, given these regions were designed as domains at the sampling stage.

Since the samples for these regions are smaller compared to the samples for the five regions (except for the Mediterranean NUTS1 region, which is the equiavalent of the South region among five regions), higher standard errors were observed for the TFR estimates. Moreover, the change in TFR after adding the second dataset was not only positive for the 12 regions. The TFRs for East Marmara, North East Anatolia and South East Anatolia slightly decreased after adding TDHS-2008 data. The largest absolute change was observed for Central East Anatolia: an increase of 16 percent.

The most interesting finding about the TFRs for NUTS 1 regions was that standard errors did not always decrease either. For Istanbul and Aegean, the standard errors increased. Other than these two regions, the gains in standard error changed between 13 percent (West Black Sea) and 35 percent (West Anatolia).

Table 4.23 Three-year total fertility rates prior to September, 2003, Turkey, place of residence

		TFR	SE(TFR)	DEFF	CV	Decrease in SE (percent)
Total	2003	2.25	0.049	1.145	0.02	
	2003 and 2008	2.38	0.042	1.230	0.02	14

Type of place of residence		TFR	SE(TFR)	DEFT	CV	Decrease in SE (percent)
Urban						
	2003	2.07	0.054	1.084	0.03	
	2003 and 2008	2.19	0.049	1.252	0.02	9
Rural						
	2003	2.68	0.114	1.381	0.04	
	2003 and 2008	2.89	0.086	1.252	0.03	25

Table 4.23 (continued). Three-year total fertility rates prior to September, 2003, Turkey, regions

Region		TFR	SE(TFR)	DEFT	CV	Decrease in SE (percent)
West						
	2003	1.89	0.061	0.569	0.03	
	2003 and 2008	2.05	0.077	1.237	0.04	-26
South						
	2003	2.32	0.168	1.896	0.07	
	2003 and 2008	2.36	0.112	1.515	0.05	33
Central						
	2003	1.90	0.095	0.943	0.05	
	2003 and 2008	2.01	0.071	0.878	0.04	25
North						
	2003	1.89	0.134	0.801	0.07	
	2003 and 2008	2.06	0.11	1.051	0.05	18
East						
	2003	3.70	0.157	2.062	0.04	
	2003 and 2008	3.91	0.109	1.573	0.03	31

Table 4.23 (continued). Three-year total fertility rates prior to September, 2003, Turkey, educational attainment

Education		TFR	SE(TFR)	DEFT	CV	Decrease in SE (percent)
No education or first level primary school incomplete						
	2003	3.71	0.159	1.471	0.04	
	2003 and 2008	4.00	0.106	0.992	0.03	33
First level primary complete, second level incomplete						
	2003	2.39	0.08	1.346	0.03	
	2003 and 2008	2.54	0.061	1.124	0.02	24
Second level and higher						
	2003	1.50	0.079	1.094	0.05	
	2003 and 2008	1.54	0.067	1.268	0.04	15

Table 4.24 Three-year total fertility rates prior to September, 2003, Turkey, 12 regions

12 Regions	TFR	SE(TFR)	DEFT	CV	Decrease in SE (percent)
Istanbul					
2003	1.86	0.089	0.771	0.05	
2003 and 2008	2.11	0.117	1.153	0.06	-31
West Marmara					
2003	1.72	0.161	0.721	0.09	
2003 and 2008	1.73	0.139	1.201	0.08	14
Aegean					
2003	1.76	0.132	0.635	0.08	
2003 and 2008	1.98	0.152	1.508	0.08	-15
East Marmara					
2003	2.11	0.146	0.375	0.07	
2003 and 2008	2.09	0.109	0.610	0.05	25
West Anatolia					
2003	1.72	0.163	1.192	0.09	
2003 and 2008	1.86	0.105	0.845	0.06	36
Mediterranean					
2003	2.32	0.168	1.896	0.07	
2003 and 2008	2.36	0.112	1.515	0.05	33
Central Anatolia					
2003	2.2	0.179	0.815	0.08	
2003 and 2008	2.37	0.136	0.895	0.06	24
West Black Sea					
2003	1.83	0.143	0.764	0.08	
2003 and 2008	1.98	0.124	1.032	0.06	13
East Black Sea					
2003	1.99	0.232	1.166	0.12	
2003 and 2008	2.24	0.177	1.065	0.08	24
North East Anatolia					
2003	3.33	0.267	1.568	0.08	
2003 and 2008	3.26	0.178	1.467	0.05	33
Central East Anatolia					
2003	3.11	0.303	2.217	0.1	
2003 and 2008	3.59	0.213	1.633	0.06	30
Southeast Anatolia					
2003	4.19	0.224	2.059	0.05	
2003 and 2008	3.92	0.153	1.672	0.04	32

The two year estimates are consistent with overlapping confidence intervals for Turkey (Table 4.23). The point estimates are higher for the West, Central and Noth regions, and are higher for the rest. The standard errors decrease with the addition of TDHS-2008 for all regions and Turkey. The one year TFRs for Turkey calculated prior to January 2004 are higher than those prior to January 2003 and 2002. For the most recent one year TFR, the lowest CV is 4 percent (East region) and the highest CV is 9 percent (North region) after combining two surveys.

Table 4.25 Two-year total fertility rates prior to January; 2004, 2002 and 2000

		January 2004				January 2003				January 2002			
		TFR	SE(TFR)	DEFF	CV	TFR	SE(TFR)	DEFF	CV	TFR	SE(TFR)	DEFF	CV
Turkey	2003	2.14	0.081	1.075	0.04	2.52	0.065	1.114	0.03	2.51	0.069	1.186	0.03
	2003 and 2008	2.12	0.063	1.073	0.03	2.68	0.056	1.177	0.02	2.54	0.053	1.129	0.02
West	2003	1.76	0.138	0.962	0.08	2.01	0.092	0.829	0.05	1.95	0.115	1.185	0.06
	2003 and 2008	1.85	0.118	1.148	0.06	2.21	0.093	1.168	0.04	1.96	0.088	1.111	0.05
South	2003	2.47	0.254	1.315	0.10	2.47	0.162	1.162	0.07	2.38	0.208	1.971	0.09
	2003 and 2008	2.17	0.158	1.127	0.07	2.62	0.109	0.888	0.04	2.55	0.145	1.536	0.06
Central	2003	1.64	0.156	0.881	0.10	2.26	0.150	1.112	0.07	2.53	0.125	0.713	0.05
	2003 and 2008	1.75	0.115	0.910	0.07	2.45	0.114	1.039	0.05	2.46	0.104	0.952	0.04
North	2003	1.56	0.198	0.914	0.13	2.43	0.173	0.880	0.07	2.31	0.199	1.208	0.09
	2003 and 2008	1.84	0.174	1.063	0.09	2.52	0.142	0.988	0.06	2.38	0.140	1.058	0.06
East	2003	3.65	0.223	1.503	0.06	4.22	0.200	1.616	0.05	4.22	0.189	1.586	0.04
	2003 and 2008	3.43	0.146	1.186	0.04	4.45	0.163	1.411	0.04	4.48	0.150	1.418	0.03

Table 4.26 Single-year total fertility rates prior to January; 2004, 2003 and 2002

		January 2004				January 2003				January 2002			
		TFR	SE(TFR)	DEFF	CV	TFR	SE(TFR)	DEFF	CV	TFR	SE(TFR)	DEFF	CV
Turkey	2003	2.14	0.08	1.075	0.04	2.13	0.09	1.061	0.04	2.51	0.09	1.104	0.04
	2003 and 2008	2.12	0.06	1.073	0.03	2.29	0.07	1.103	0.03	2.74	0.08	1.152	0.03
West	2003	1.76	0.14	0.962	0.08	1.95	0.14	0.836	0.07	1.97	0.13	0.806	0.07
	2003 and 2008	1.85	0.12	1.148	0.06	1.93	0.12	1.078	0.06	2.38	0.14	1.129	0.06
South	2003	2.47	0.25	1.315	0.10	2.08	0.23	1.250	0.11	2.43	0.27	1.434	0.11
	2003 and 2008	2.17	0.16	1.127	0.07	2.34	0.18	1.165	0.08	2.58	0.18	1.125	0.07
Central	2003	1.64	0.16	0.881	0.10	1.85	0.19	1.101	0.10	2.26	0.22	1.090	0.10
	2003 and 2008	1.75	0.12	0.910	0.07	1.92	0.14	1.090	0.07	2.40	0.17	1.067	0.07
North	2003	1.56	0.20	0.914	0.13	1.60	0.20	0.642	0.13	2.52	0.24	0.720	0.10
	2003 and 2008	1.84	0.17	1.063	0.09	1.76	0.16	0.855	0.09	2.64	0.19	0.810	0.07
East	2003	3.65	0.22	1.503	0.06	3.20	0.17	1.031	0.05	4.32	0.28	1.493	0.07
	2003 and 2008	3.43	0.15	1.186	0.04	3.94	0.16	1.116	0.04	4.46	0.22	1.222	0.05

4.2.2 Infant Mortality Rate

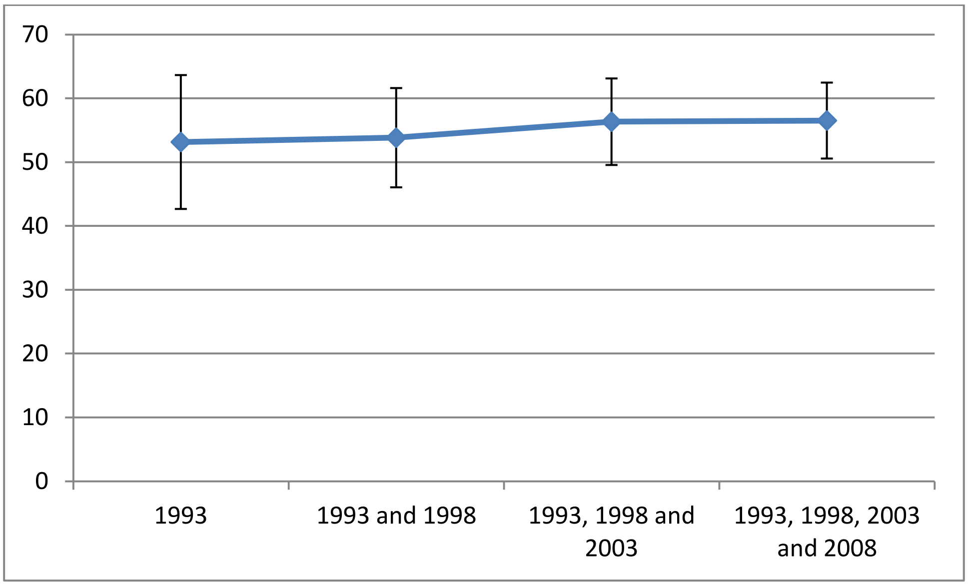
The IMR is usually calculated for a 5 year period of time when reporting for the whole country in the Turkish DHSs. This is has also been the case for basic characteristics such as type of place of residence, region and mother's education, until TDHS-2013, where the time period for these estimates was increased to 10 years. The reason for this was the decrease in the number of infant deaths, it became a rare event in time. As infant deaths got more rare, the time period of rate estimates were extended so that more cases would be captured; and standard errors would be lower. This section re-calculates the IMRs with the addition of extra data sets. Just like the TFR, the addition of extra data sets was done for the overall sample and some basic variables: Type of place of residence, region and education.

TDHS-1993

Just as done for the TFR, an artificial date was selected to be used as interview date in order to compute IMRs corresponding to a five year period prior to TDHS-1993 (September 1993) from later datasets. Therefore births recorded up to this date were included in the calculation of the rate, and not the rest.

The IMR for Turkey goes up slightly with each additional dataset, however, this increase is not significant. All confidence intervals overlap (Figure 4.6). The IMR is estimated at 53 per 1000 from TDHS-1993, and at 57 per 1000 from all four data sets; which is a very small change considering the rareness of the event. It should be noted that the confidence interval is much lower as more datasets are added; the coefficient of variation drops from 0.10 to 0.05 in the meantime (Table 4.26). The decrease in the standard error is 43 percent.

Figure 4.6 Five-year infant mortality rates (per 1000) prior to September, 1993 and their standard errors



A pattern similar to the overall estimate is observed for urban and rural areas. The estimates undergo minor changes with the additional data sets, with standard errors – thus coefficient of variations – almost halved. The changes in the point estimates are around 10 percent, yet the decreases in the standard errors are around 40-45 percent after three additional datasets.

For the five demographic regions, the IMR estimates show increases with additional datasets for the West, North and East regions. The opposite is observed for South and Central regions. Again, all confidence intervals overlap for all these estimates.

The estimates of IMR by mother's educational status show increasing patterns of IMR for all three groups within overlapping confidence intervals. Due to the low number of cases in the secondary level primary school education and higher category, the coefficient of variation is very high to begin with (39 percent), which decreases to 16 percent after the other data sets have been added.

Table 4.27 Five-year infant mortality rates (per 1000) prior to September, 1993, Turkey, place of residence

		IMR	SE(IMR)	DEFF	CV	Decrease in SE (percent)
Total	1993	53.16	3.850	1.107	0.07	
	1993 and 1998	53.84	2.792	1.145	0.05	27
	1993, 1998 and 2003	56.34	2.519	1.409	0.05	35
	1993, 1998, 2003 and 2008	56.52	2.182	1.355	0.04	43

Type of place of residence		IMR	SE(IMR)	DEFF	CV	
Urban						
	1993	44.85	4.795	1.210	0.11	
	1993 and 1998	47.36	3.403	1.192	0.07	29
	1993, 1998 and 2003	49.67	3.141	1.583	0.06	34
	1993, 1998, 2003 and 2008	47.89	2.625	1.493	0.06	45
Rural						
	1993	65.64	6.201	0.947	0.09	
	1993 and 1998	64.16	4.424	0.922	0.07	29
	1993, 1998 and 2003	67.45	4.042	1.103	0.06	35
	1993, 1998, 2003 and 2008	72.34	3.724	1.092	0.05	40

Table 4.27 (continued). Five-year infant mortality rates (per 1000) prior to September, 1993, regions

Region		IMR	SE(IMR)	DEFF	CV	Decrease in SE (percent)
West						
	1993	43.07	8.092	1.302	0.19	
	1993 and 1998	47.93	6.119	1.268	0.13	24
	1993, 1998 and 2003	51.74	5.914	1.764	0.11	27
	1993, 1998, 2003 and 2008	50.47	4.816	1.523	0.10	40
South						
	1993	55.23	6.471	0.632	0.12	
	1993 and 1998	52.24	5.018	0.760	0.10	22
	1993, 1998 and 2003	47.15	4.251	0.889	0.09	34
	1993, 1998, 2003 and 2008	47.06	3.836	0.914	0.08	41
Central						
	1993	59.95	8.065	0.933	0.14	
	1993 and 1998	54.29	5.687	0.920	0.11	29
	1993, 1998 and 2003	54.03	5.322	1.153	0.10	34
	1993, 1998, 2003 and 2008	52.54	4.562	1.100	0.09	43
North						
	1993	46.82	10.075	1.341	0.22	
	1993 and 1998	55.99	7.914	1.272	0.14	21
	1993, 1998 and 2003	56.04	6.071	1.063	0.11	40
	1993, 1998, 2003 and 2008	58.19	6.148	1.286	0.11	39
East						
	1993	59.23	8.449	0.966	0.14	
	1993 and 1998	60.38	5.651	1.071	0.09	33
	1993, 1998 and 2003	69.44	4.338	1.022	0.06	49
	1993, 1998, 2003 and 2008	73.13	3.854	1.028	0.05	54

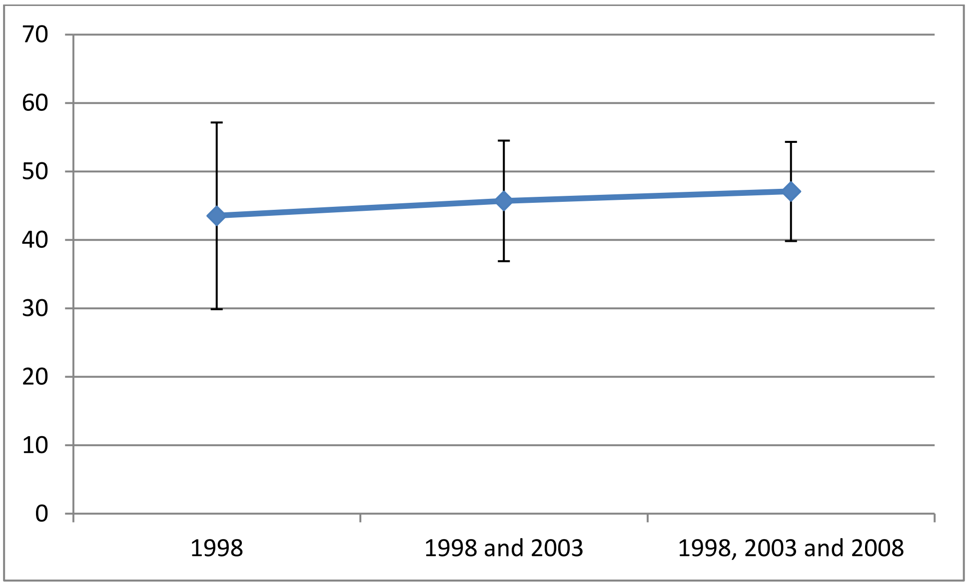
Table 4.27 (continued. Five-year infant mortality rates (per 1000) prior to September, 1993, educational attainment of mother

Education	IMR	SE(IMR)	DEFF	CV	Decrease in SE (percent)
No education or first level primary school incomplete					
1993	57.91	5.577	0.752	0.10	
1993 and 1998	58.60	4.055	0.835	0.07	27
1993, 1998 and 2003	65.80	3.333	0.834	0.05	40
1993, 1998, 2003 and 2008	70.21	3.149	0.910	0.05	44
First level primary complete, second level incomplete					
1993	52.47	5.144	1.028	0.10	
1993 and 1998	55.30	3.826	1.036	0.07	26
1993, 1998 and 2003	55.21	3.296	1.188	0.06	36
1993, 1998, 2003 and 2008	53.85	2.797	1.119	0.05	46
Second level+					
1993	33.97	13.065	2.631	0.39	
1993 and 1998	33.46	6.654	1.329	0.20	49
1993, 1998 and 2003	36.98	8.188	2.829	0.09	-23
1993, 1998, 2003 and 2008	37.28	6.086	1.949	0.16	26

TDHS-1998

The artificial date for the computation of IMRs corresponding to a five year period prior to TDHS-1998 from later datasets was set as September, 1998. The point estimates for the overall sample are quite similar for the TDHS-1993 and combined data set. The IMR increased by 8 percent after all data sets were used, and the standard error reduced by 47 percent (Table 4.28).

Figure 4.7 Five-year infant mortality rates (per 1000) prior to September, 1998 and their standard errors



For urban and rural areas the IMR estimates increased by 11 percent and by 9 percent respectively. Neither were significant changes. The decreases in standard errors were 33 and 46 respectively, with coefficients of variation almost halved.

Regional IMR estimates were subjected to changes between -13 percent and 17 percent. The negative percent change was observed for the North region. The lowest gain of standard errors was also observed for this region: It was reduced by 17 percent. The highest increase was observed for the South region. The largest gain in the standard error is obtained for the East region, with the coefficient of variation dropping from 20 percent to 8 percent.

For the estimates of infant births of the lower two educational categories the point estimates increased by 26 percent and 11 percent respectively, with standard errors lowered by 32 and 50 percent respectively. For the more educated category, the combined data set estimate decreased with the addition of TDHS-

2003, and stayed close to that after the addition of TDHS-2008. The decrease in the coefficient of variation is relatively small for this category.

Table 4.28 Five-year infant mortality rates (per 1000) prior to September, 1998, Turkey, place of residence

		IMR	SE(IMR)	DEFF	CV	Decrease in SE (percent)
Total	1993	43.52	4.834	2.016	0.11	
	1998 and 2003	45.70	3.139	1.885	0.07	35
	1998, 2003 and 2008	47.08	2.577	1.817	0.06	47
Type of place of residence						
Urban						
	1998	36.16	4.136	1.130	0.11	
	1998 and 2003	37.15	3.114	1.501	0.08	25
	1998, 2003 and 2008	40.26	2.756	1.621	0.07	33
Rural						
	1998	55.62	9.081	2.025	0.16	
	1998 and 2003	61.11	6.016	1.769	0.10	34
	1998, 2003 and 2008	60.69	4.897	1.700	0.08	46

Table 4.28 (continued) Five-year infant mortality rates (per 1000) prior to September, 1998, regions

Region		IMR	SE(IMR)	DEFF	CV	Decrease in SE (percent)
West						
	1998	33.83	8.276	1.459	0.25	
	1998 and 2003	31.30	5.495	1.706	0.18	34
	1998, 2003 and 2008	35.71	4.788	1.638	0.13	42
South						
	1998	32.77	8.605	1.578	0.26	
	1998 and 2003	34.14	5.214	1.169	0.15	39
	1998, 2003 and 2008	36.85	4.597	1.252	0.18	47
Central						
	1998	41.80	8.554	1.272	0.21	
	1998 and 2003	40.34	6.506	1.488	0.16	24
	1998, 2003 and 2008	47.42	5.492	1.323	0.12	36
North						
	1998	44.91	8.334	0.760	0.19	
	1998 and 2003	46.92	8.790	1.530	0.19	-5
	1998, 2003 and 2008	39.72	6.929	1.575	0.17	17
East						
	1998	62.16	12.567	2.846	0.20	
	1998 and 2003	72.96	6.791	2.016	0.09	46
	1998, 2003 and 2008	69.34	5.243	1.907	0.08	58

Table 4.28 (continued) Five-year infant mortality rates (per 1000) prior to September, 1998, educational attainment of mother

Education		IMR	SE(IMR)	DEFF	CV	Decrease in SE (percent)
No education or first level primary school incomplete						
	1998	48.21	5.629	0.723	0.12	
	1998 and 2003	58.39	4.449	1.020	0.08	21
	1998, 2003 and 2008	60.61	3.801	1.088	0.06	32
First level primary complete, second level incomplete						
	1998	41.10	7.054	2.418	0.17	
	1998 and 2003	43.17	4.240	1.812	0.10	40
	1998, 2003 and 2008	45.59	3.524	1.748	0.08	50
Second level and higher						
	1998	41.18	6.315	0.637	0.15	
	1998 and 2003	28.38	4.252	0.878	0.15	33
	1998, 2003 and 2008	29.15	3.844	0.976	0.13	39

TDHS-2003

The artificial date of interview was fixed at January 2004 for the IMR estimates of 2003 from multiple datasets. The 5 year IMR estimate for Turkey was computed as 29 per 1000 from TDHS-2003, and it was found as 31 per thousand after the addition of TDHS-2008. The coefficient of variation decreased from 11 percent to 7 percent.

Among regional estimates, West and North yielded similar estimates with the combined data set and TDHS-2003 data set. The other three showed increases between 10-20 percent. While the decreases in standard errors led to decreases in the coefficient of variation, they remained at 19 percent and higher for all regions except the East.

Within different educationa levels, the two lower categories of education have smaller coefficients of variation, because they are based on more women. The point estimates for all three categories slightly increase, and the standard errors are reduced around one third for the first two categories. There is no gain in standard errors for the IMR estimates from women who have and educational level of high school or higher.

Figure 4.8 Five-year infant mortality rates (per 1000) prior to January, 2004 and their standard errors

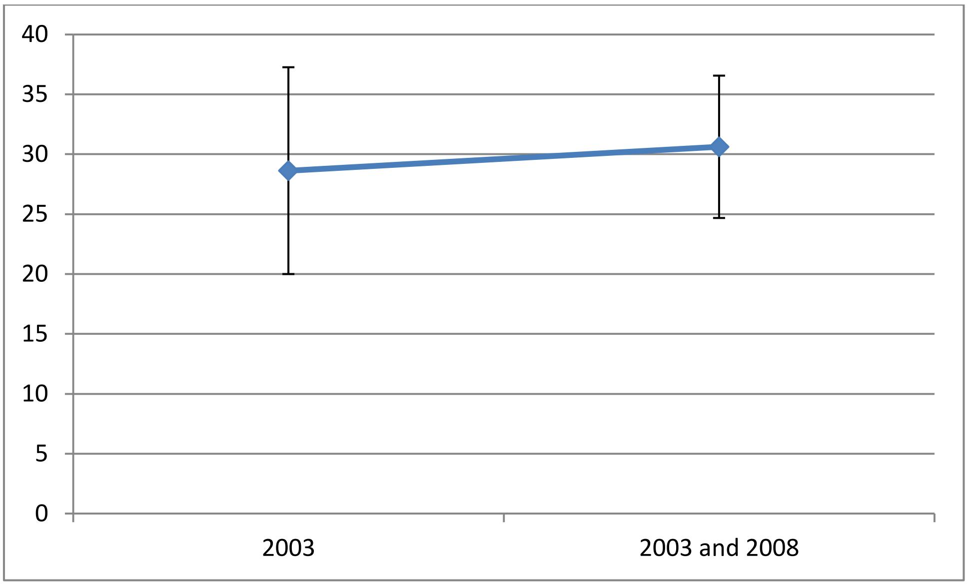


Table 4.29 Five-year infant mortality rates (per 1000) prior to January, 2004, Turkey, place of residence

		IMR	SE(IMR)	DEFF	CV	Decrease in SE (percent)
Total	2003	28.62	3.15	1.64	0.11	
	2003 and 2008	30.61	2.14	1.36	0.07	32

Type of place of residence		IMR	SE(IMR)	DEFF	CV	
Urban						
	2003	23.78	3.07	1.27	0.13	
	2003 and 2008	24.96	2.28	1.27	0.09	26
Rural						
	2003	38.16	7.02	1.96	0.18	
	2003 and 2008	42.94	4.63	1.50	0.11	34

Table 4.29 (contiued. Five-year infant mortality rates (per 1000) prior to January, 2004, regions

Region		IMR	SE(IMR)	DEFF	CV	Decrease in SE (percent)
West						
	2003	22.35	6.31	1.81	0.28	
	2003 and 2008	21.56	4.16	1.44	0.19	34
South						
	2003	26.74	7.85	1.43	0.29	
	2003 and 2008	31.94	6.02	1.32	0.19	23
Central						
	2003	20.55	6.19	1.23	0.30	
	2003 and 2008	23.17	4.33	1.11	0.19	30
North						
	26.741	34.32	12.28	1.71	0.36	
	31.943	32.40	8.47	1.80	0.26	31
East						
	26.741	41.42	5.36	1.42	0.13	
	31.943	46.72	3.55	1.07	0.08	34

Table 4.29 (continued). Five-year infant mortality rates prior to January, 2004, educational attainment of mother

Education	IMR	SE(IMR)	DEFF	CV	Decrease in SE (percent)
No education or first level primary school incomplete					
2003	36.71	4.78	0.97	0.13	
2003 and 2008	40.12	3.16	0.75	0.08	34
First level primary complete, second level incomplete					
2003	28.75	4.06	1.28	0.14	
2003 and 2008	30.99	2.86	1.14	0.09	30
Second level and higher					
2003	13.75	2.77	0.54	0.20	
2003 and 2008	16.28	2.74	0.81	0.17	1

5. DISCUSSION

The aim of this dissertation was to combine data sets from Turkey Demographic and Health Surveys with a statistically correct methodology and use the combined data sets to re-calculate widely used demographic indicators. The main hypothesis was that combining these data would produce statistics with higher precision because the number of cases for analysis would increase. New sampling weights were calculated and new stratification variables were created for this purpose, and SAS Macros were modified to get correct estimates of standard errors of rates.

The SAS macros were originally obtained from Thiam and Aliaga (2001) upon request. The SAS macros required time and computer effort to run. Especially for the simple random sample standard errors, the Jackknife loop had to run as many times as 25,000 for the combined women’s data set, and even more for the combined birth histories data set. Going over the SAS macros step by step led to a deeper understanding of the variance estimation methods used by DHS. This in turn led to the selection of a JRR n method rather than a JRR 1 method in this thesis, as opposed to the DHS choice of Jackknife 1, which takes stratification into account. This was preferred both to be consistent with the stratification used by DHS for Taylor series estimations of variances, and because it would be statistically correct not to ignore stratification.

There is a vast literature on combining surveys, and the closest examples were the integrated survey series mostly found in the US. Most examples with weights that reflect population sizes recommended using original weights. With TDHS data, the weights were first brought to population scales, and then they were re-scaled to match the total sample size of interviewed women to be in line with the DHS approach of weight scaling. Methodological documentation from DHS that was published shortly after the computations for this thesis were made suggested a similar approach, confirming the approach taken up in this dissertation.

Combining data for time invariant variables that are little affected by various survey errors (such as selection bias) was the most useful among the applications. Means and proportions that were expected to change over time provided estimates over a period of five years, and would be average values.

For all types of indicators, standard errors decreased with additional surveys. However, the gains in standard error got less with each additional survey. This is because the square of standard error (sampling variance) is inversely proportional with sample size. In other words, if the sample size is doubled, the sampling variances decreases by half; but being its square root, the standard error decreases by 41 percent. If the sample size is quadrupled, the standard error decreases by 50 percent. Because of this relationship; theoretically; the addition of the first survey is expected to yield the largest gain in precision. In practice, however, the decreases in standard error were less than this theoretical expectation, as were seen in the Results section. Yet the percent decreases were less after two surveys were added to an initial data set, a phenomenon that could only be observed for TDHS-1993. A fourth survey brings less improvement in terms of precision.

Educational attainment was selected as a proportion that was expected to remain stable over time. Since three of the TDHSs interviewed ever married women only; cohort estimates of educational attainment was problematic for young cohorts of women. The marital composition of young cohorts changed drastically over time; and given the correlation of marital age and education, so did estimates for education. For older cohorts, estimates were stable, and there were gains in precision.

The mean number of children was defined for up to fixed dates in time, so that it would not vary over time for cohorts, which also produced fairly consistent estimates as a result of combining more data. Only for TDHS-1993 combining surveys increased the estimates to some extent, which could be related to underreporting in this survey. Standard errors were low to begin with for most birth cohorts, and they reduced further with more data sets. The standard error for the overall estimate of mean number of children ever born decreased by only very little, for which extra weighting was used, which might have caused an inflation in the standard errors.

Ideal number of children and contraceptive use provided cross-sectional information that could and would change at each survey date, thus the combined estimates of these were less meaningful, providing averages over a 5 year interval. This would have little value for policy making purposes for Turkey as a whole, because their levels of precision are high enough in single surveys (for example, coefficients of variations were less than 5% for all major types of contraceptives but folkloric methods). They could be of more use for smaller domains if required.

The calculated TFR for two or single years showed that combining data sets can be useful in terms of statistical precision of shorter duration rates. The coefficients of variation observed after combining data from all four surveys for the one year TFR were close to those obtained from a single survey. For the calculations of TFR and its standard error for the twelve statistical regions, there were some unexpected findings. For some regions (Istanbul and the Aegean regions), the standard errors increased after combining data from TDHS-2003 and TDHS-2008. It can be argued here that although standard errors are related to sample size, they also depend on the dispersion of the sample as well. Heterogeneity within these regions might have increased over a period of 5 years. Especially keeping in mind that Western regions as such are popular origins for international migration, the population composition of these regions are prone to change in time.

Because of its rare event property, infant mortality has been the demographic indicator for which combining surveys was most beneficial. For example, the coefficient of variation decreased to 4 percent when four surveys are combined for the five years preceding September 1993, compared to a coefficient of variation of 14 percent from TDHS-1993 only. For the Western region in TDHS-1993, an IMR of 43 per thousand was estimated with a standard error of 6.5, meaning 95 percent of the samples that are designed the same would yield IMRs between 21 and 66. The confidence interval for the corresponding IMR estimate was almost half as narrow: between 38 and 63.

The total fertility rate presented an interesting case: Going back in time through birth histories for the date of an earlier survey mostly resulted in higher estimates of fertility than what the earlier survey suggested. This phenomenon previously has been pointed out in the literature (Arnold and Blanc, 1990; Curtis and Arnold, 1994; Komba and Aboud, 1994; Pison *et al.*, 1995; Hancıoğlu, 1997; Blanc and Gray, 2000; Rutstein, 2002; Coşkun, 2008). The long mother and child health module for children under 5 years of age was claimed to be related to cheating by interviewers, through shifting of births to earlier years so as to miss this section, or leave them out completely. Previous work by the author as regards data quality (not shown) was in line with this claim: Five year TFRs were calculated from TDHS-1993, TDHS-1998, TDHS-2003 and TDHS-2008 for the 0-4, 5-9 and 10-14 years preceding survey date to find out how they agree between different surveys. Without exceptions, the 0-4 year TFR estimate of any survey was lower than the 5-9 year TFR estimate from the next survey. The differences were 0.51, 0.18 and 0.27 respectively between TDHS-1998 and TDHS-1993, TDHS-2003 and TDHS-1998, and TDHS-2008 and TDHS-2003.

The following quotation includes Hancıoğlu's (1997) conclusions based on his study on TDHS-1993 and earlier surveys: "*Comparing rates obtained from the four surveys reveals the problematic nature of birth history data in Turkey… Rates obtained for five-year periods immediately preceding each survey are, as mentioned earlier, too low on the basis of trends implied by estimates for earlier periods, and more importantly, based on estimates obtained for the same periods from surveys carried out later…Typically, such curves are indicative of omissions* *of early births by older women, of omissions and/or misplacement of recent births into the period 5-9 years preceding the surveys…The consequence is an overstatement of recent declines in fertility, while fertility rates in the distant past are underestimated… It has become clear that accepting the results of Turkish demographic surveys at face value would generally lead to misleading interpretations of fertility levels and trends…*". It should be underlined here that this study was completed before three of the surveys used in this thesis were even carried out; however, it is important because it provides an insight to the problem faced in this thesis. The pattern mentioned by Hancıoğlu (1997) was observed in the findings of this thesis as well. Therefore combining surveys for TFR caused point estimates to usually increase. This problem was observed most apparently for TDHS-1993, and it seems to be the least of a problem in TDHS-1998. TDHS-2003 was somewhere in between these two in terms of this problem. In most cases, although changes in point estimates were observed, they had overlapping confidence intervals.

The estimates for the IMR were more consistent than TFR. There were slight increases in the point estimates going back in time, yet they were not as apparent as those for the TFR. Birth year ratios (not shown) in the breakdown of survival status did not show major differences. It may thus be speculated that there were children who were reported earlier than they were born as suggested above related to TFR, but this did not introduce bias in terms of mortality.

6. CONCLUSION

This thesis aimed to combine data across four Turkey Demographic and Health Surveys (TDHS) in a methodologically correct way, to estimate a variety of demographic indicators including proportions, means and rates. The hypothesis was that combining data for this purpose would provide estimates with higher statistical precision, which could become useful for smaller domains, rare characteristics / events, and shorter durations (for rates).

Calculations were made by using the SAS software. Built in procedures were used for simple statistics (means and proportions except for the mean number of children ever born), and SAS macros were edited and used for complex proportions and demographic rates that used the Jackknife repeated replications method for complex sample variance estimation.

The findings for each selected statistic led to different conclusions depending on the nature of the statistic; except for the general theme of decreasing standard errors (square root of sampling variances). Gains in precision were observed for all types of variables with the addition of more data. For means and proportions that are available with a fair level of precision, combining data would prove to be useful for small domains.

Total fertility rate and infant mortality rate are among the most important outputs of DHSs in Turkey. Since vital registration data is subject to coverage and late reporting issues, and censuses are based on a small set of questions, surveys have been the optimal data source for the direct estimation of these rates in Turkey since the 1960s. Thus demographic rates have been paid extra attention for this dissertation. Combining for the total fertility rate raised questions on data quality that is both documented in the literature, and observed in this thesis. For IMR, given its rare event property and large confidence intervals, combined THDS data seems to work well.

The challenges included a weighting adjustment for the combined data sets, calculating entirely new weights for the estimation of some selected demographic indicators, and computing the complex sample and simple random sample standard errors of the total fertility rate and infant mortality rate. The different samples were treated as independent. After standardizing and pooling data into a single data file, the sample weights were adjusted via denormalization (as called by ICF International, 2012) and rescaling, according to the population figures of Turkey at the survey dates. Equal weighting of surveys, still using the sample weights was employed for some cohort estimates, and weighting according to the original survey’s age structure was preferred for some age related demographic estimates from combined surveys.

The contribution of this thesis to the literature was the suggestion of a statistically correct approach for combining TDHSs across time, provide a different than usual exploitation of widely used TDHS data and provide higher precision demographic indicators for Turkey.

A number of limitations need to be considered about this dissertation. Firstly, while the individual questionnaires of three of the surveys were based on ever married women only, one was based on all women. Thus indicators, the denominators of which were based on all women were subjected to different methodologies (mean number of children ever born and total fertility rate). The potential effect of this was not further analyzed so as to keep a standard with the TDHS main reports. Secondly, the data quality problem mentioned in the discussion section that is related to interviewer error needs to be studied further so as to tackle it in future demographic surveys. Thirdly, types of statistics that were not presented at the Appendices of DHS surveys were left out in this study: Further work should be done to see the implications of using an aggregated data source on the standard errors of different statistics, such as medians and odds ratios.

REFERENCES

Aldrich, J. (2008). *Professor A. L. Bowley's Theory of the Representative Method*. Sample Surveys and Bayesian Statistics Conference, Southampton, UK.

Arnold, F. and Blanc, A. K. (1990). *Fertility Levels and Trends*. Institute for Resource Development/Macro Systems, Inc., DHS Comparitive Studies.

Bethelem, J. (2009). *The Rise of Survey Sampling*. Statistics Netherlands Discussion Paper, 09015.

Bjerve, P. J. (1973). *World Fertility Survey*. Arbeidsnotater.

Blanc, A. K. and Gray, S. (2000). *Greater Than Expected Fertility Decline in Ghana: An Examination of the Evidence*. Macro International Inc and National Population Council Secretariat.

Brainard, J. and Zaharlick, A. (1992). *Lifestyle Change and Fertility Transition for Southeast Asian Refugeees: Focus on the Lao*. University of Pennsylvania, MASCA Research Papers in Science and Archaeology.

Brick, J. M., Morganstein, D. and Valliant, R. (2000). *Analysis of Complex Sample Data Using Replication*. Westat.

Brunt, L. (2001). *The Advent of the Sample Survey in the Social Sciences*. Journal of the Royal Statistical Society, 50, (2), 179-189.

Center for Diesase Control. (2013). *Variance Estimation and Other Analytic Issues, NHIS 2006-2012 (Adapted from the 2006-2012 NHIS Survey Description Documents - appendices "Variance Estimation and Other Analytic Issues", and "Merging Data Files and Combining Years of Data")* http://www.cdc.gov/nchs/data/nhis/2006var.pdf, Retrieved February 1, 2014.

Chand, N. and Alexander, C. H. (2000). *Multi-year averages from a rolling sample survey*. American Statistical Association Annual Preceedings

Chu, A. and Goldman, J. (1997). *Weighting procedures for USDA's continuing survey of food intakes by individuals 1994-96*. American Statistical Association, Section on Survey Research Methods, Anaheim, California.

Coşkun, Y. (2008). *Data Quality Assessment of Birth History Data in Turkey Demographic and Health Surveys* Institute of Population Studies Technical Demography Program, Unpublished PhD Dissertation.

Curtis, S. L. and Arnold, F. (1994). *An Evaluation of the Pakistan DHS Survey Based on the Reinterview Survey*. Occasional Papers.

DHS Program The DHS Program. (2014). *Where we work / Country List*, http://dhsprogram.com/Where-We-Work/Country-List.cfm, Retrieved March 12, 2014.

Dong, Q. (2012). *Combining Information from Multiple Complex Surveys*. Survey Methodology Programme, Unpublished PhD Dissertation.

Duncan, W. G. (1973). *Fertility and Related Surveys*. World Fertility Survey (WFS) Occasional Paper, 1.

Durbin, J. (1959). *A note on the application of Quenoille's method of bias reduction to the estimation of ratios*. Biometrika, 60, 551-558.

Ericksen, E. P. (1973). *A Method for Combining Sample Survey Data and Symptomatic Indicators to Obtain Population Estimates for Local Areas*. Demography, 10, (2), 137-160.

Fay, R. E. and Herriot, R. A. (1979). *Estimates of income for small places: an application of James-Stein procedures to census data*. Journal of the American Statistical Association, 269-277.

Fienberg, S. E. and Tanur, J. M. (1966). *Reconsidering the Fundamental Contributions of Fisher and Neyman on Experimentation and Sampling*. International Statistical Review, 64, (3), 237-253.

Frankel, M. R. (1971). *Inference from survey samples: An empirical investigation*, Institute for Social Research, University of Michigan, Ann Arbor, Michigan.

Friedman, E. M. and Jang, D. (2002). *Combined Estimates from Four Quarterly Survey Data Sets*. Joint Statistical Meetings Section on Survey Research Methods.

Halfon, S. (2007). *The Cairo Consensus: Demographics Surveys, Women's Empowerment, and Regime Change in Population Policy*, Lexington Books, Lanham.

Hancıoğlu, A. (1997). *Fertility Trends in Turkey: 1978-1993*, Ankara, Calverton.

Hansen, M. H. (1987). *Some History and Reminesces on Survey Sampling*. Statistical Science, 2, (2), 180-190.

Hassett, N. C. and Saavedra, P. J. (1995). *Combining data to produce timely estimates*. American Statistical Association Joint Statistical Meetings, Orlando, Florida.

HUIPS. (1994). *Turkey Demographic and Health Survey 1993*. Hacettepe University Institute of Population Studies, Ministry of Health Directorate of Morther and Child Health, Demographic and Health Surveys, Macro International Int.

HUIPS. (1999). *Turkey Demographic and Health Survey 1998*. Hacettepe University Institute of Population Studies, Macro International Inc., Ministry of Health Directorate of Mother and Child Health, Nuited Nations Population Fund, United States Agency for International Development.

HUIPS. (2005). *Turkey Demographic and Health Survey 2003*. Hacettepe University Institute of Population Studies, Ministry of Health Directorate of Mother and Child Health, T.R Prime Ministry State Organization of Planning, European Union.

HUIPS. (2009). *Turkey Demographic and Health Survey 2008*. Hacettepe University Institute of Population Studies, Ministry of Health Directorate of Mother and Child Health and Family Planning, State Planning Organization.

ICF International. (2012). *Sampling and Household Listing Manual*.

Jang, D. and Sukasih, A. (2004). *Replicate Variance Estimation For Combined Data of Independent Surveys*. Joint Statistical Meetings, Toronto.

Karatepe, Z. S. (2010). *The Relationship between Socioeconomic Differences and Economic Uncertainty: Analysis of Infant and Child Mortality in Turkey*. Department of Economic History, Unpublished Master's Dissertation.

Kırdar, M. G., Dayıoğlu, M. and Koç, İ. (2012). *The Effects of Compulsory Schooling Laws on Teenage Marriage and Births in Turkey*. Munich Personal RePEc Archive.

Kish, L. (1965). *Survey Sampling*, John Wiley and Sons Inc., New York.

Kish, L. (1979a). *Samples and Censuses*. International Statistical Review, 47, (2), 99-109.

Kish, L. (1979b). *Rolling Samples instead of Censuses*. Asian and Pacific Census Forum, G(1), 1-2, 12-13.

Kish, L. (1981a). *Using Cumulated Rolling Samples to Integrate Census and Survey Operations of the Census Bureau*, U.S. Government Printing Office, Washington, D.C.

Kish, L. (1983). *Data Collection for Details over Space and Time*, Academic, New York.

Kish, L. (1987). *Statistical Research Design*, John Wiley & Sons, New York.

Kish, L. (1990). *Rolling Samples and Censuses*. Survey Methodology, 16, (63-79).

Kish, L. (1992). *Weighting for Unequal Pi*. Journal of Official Statistics, 8, (2), 183-200.

Kish, L. (1997). *Periodic and Rolling Samples and Censuses*, Clarendon Press, Oxford.

Kish, L. (1998). *Space/Time Variations and Rolling Samples*. Journal of Official Statistics, 14, (1), 31-46.

Kish, L. (1999). *Cumulating/combining population surveys*. Survey Methodology, 25, (2), 129-138.

Kish, L. (2000). *New paradigms (models) for probability sampling*. Survey Methodology, 28, (1), 96-102.

Kish, L. and Verma, V. (1983). *Census plus Samples: Combined Uses and Designs*. Bulletin of the International Statistical Institute, 50, (1), 66-82.

Kish, L. and Verma, V. (1986). *Complete Censuses and Surveys*. Journal of Official Statistics, 2, (4), 381-395.

Koç, İ. and Eryurt, M. A. (2013). *Türkiye'de Akraba Evlilikleri ile Bebek Ölümleri Arasındaki Nedensel İlişkinin Analizi* TÜBİTAK Project No: 111K335.

Koç, İ. and Ünalan, T. (2000). *Adolescent Reproductive Behaviour in Turkey*. Turkish Journal of Population Studies, 22, 37-55.

Komba, A. S. and Aboud, S. M. (1994). *Ferility Levels, Trends, and Socioeconomic Differentials: Findings from the Tanzania Demographic and Health Survey*. DHS Regional Analysis Workshop for Anglophone African.

Lee, E. S. and Forthofer, R. N. (2006). *Strategies for Variance Estimation,* SAGE.

Lewis, G. L. (1983). *The Contraceptive Prevalence Survey Project: Content and Status* Population Index, 49, (2189-198).

Macro International Inc. . (1996). *Sampling Manual*. Macro International Inc., DHS-III Basic Documentation.

Mariott, B. M. and others. (2007). *Preliminary Data for Demographic and Health Surveys on Infant Feeding in 20 Developing Countries*. The Journal of Nutrition, 137, 518-523.

Merkouris, T. (2004). *Combining Independent Regression Estimators from Multiple Surveys*. Journal of the American Statistical Association, 99, (468), 1131-1139.

Merkouris, T. (2006). *Efficient Small-Domain Estimation by Combining Information from Multiple Surveys through Regression*. ASA Section on Survey Research Methods.

Merkouris, T. (2010). *Combining information from multiple surveys by using regression for efficient small domain estimation*. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 72, (1), 27-48.

Michaud, S. and others. (1995). *Combining administrative and survey data to reduce respondent burden in longitudinal surveys*. American Statistical Association Annual Meetings, Orlando, Florida.

Minnesota Population Center. (2012). *Integrated Public Use Microdata Series Web Page*, http://usa.ipums.org/usa/index.shtml, Retrieved August 23, 2012.

Minnesota Population Center and State Health Access Data Assistance Center. (2012). *Integrated Health Interview Series Web Page*, http://www.ihis.us/ihis/index.shtml, Retrieved August 23, 2012.

National Science Foundation. (2012). *Scientists and Engineers Statistical Data System*, http://www.nsf.gov/statistics/sestat/start.cfm, Retrieved August 26, 2012.

Nusser, S. M. and others. (2004). *Combining Data From State and National Monitoring Surveys to Assess Large-Scale Impacts of Agricultural Policy*. Journal of Agricultural, Biological, and Environmental Statistics, 9, (3), 1-17.

NVİ (İçişleri Bakanlığı Nüfus ve Vatandaşlık İşleri Genel Müdürlüğü, General Directory of Civil Registration and Nationality). (2014). *Adres Kayıt Sistemi*, http://www.nvi.gov.tr/Hakkimizda/Projeler,Aks.html, Retrieved April 20, 2014.

O'Muircheartaigh, C. and Pedlow, S. (2002). *Combining Samples vs. Cumulating cases: A Comparison of two Weighting Strategies in NLSY97*. Joint Statistical Meetings, New York City, New York.

Pison, G. and others. (1995). *Population dynamics of Senegal*, National Academies Press, Washington D.C.

Pullum, T. W. (2008). *An Assessment of the Quality of Data on Health and Nutrition in the DHS Surveys, 1993-2003* DHS Methodological Reports, 6.

Qian, J. and Kaplan, B. (2001). *Analysis of Design Effects for NAEP Combined Samples*. Annual Meeting of the American Statistical Association, Atlanta, Georgia.

Qian, J., Kaplan, B. and Weng, V. (2002). *Statistical Power Analysis for NAEP Combined Ntional and State Samples*. Joint Statistical Meetings, Section on Survey Research Methods, New York City, New York.

Raghunathan, T. E. (2006). *Combining information from multiple surveys for assessing health disparities*. Allgemeines Statistisches Archiv, 90, (4), 515-526.

Raghunathan, T. E. and others. (2007). *Combining information from two surveys to estimate county-level prevalence rates of cancer risk factors and screening*. Journal of the American Statistical Association, 102, (478), 474-486.

Rao, C. R. (1973). *Prasantha Chandra Mahalanobis. 1893-1972*. Biographical Memoirs of Fellows of the Royal Society, 19, 454-492.

Rendall, M. S. and others. (2006). *Combining Population and Survey Data for Male Fertility Estimation*. Transition to Fatherhood Working Papers, 2006-2.

Renssen, R. H. and Nieuwenbroek, N. J. (1997). *Aligning Estimates for Common Variables in Two or More Sample Surveys*. Journal of the American Statistical Association, 92, (437), 368-374.

Retherford, R. and others. (2010). *Multivariate Analysis of Parity Progression–Based Measures of the Total Fertility Rate and Its Components*. Demography, 47, (197-124).

Rosegrant, S. (2012). *ISR and the Truman/Dewey upset*. ISR Sampler.

Rutstein, S. O. (2002). *Fertility Levels, Trends, and Differentials 1995-1999*. ORC Macro, DHS Comparitive Reports.

Rutstein, S. O. (2008). *Further Evidence of the Effects of Preceding Birth Intervals on Neonatal, Infant, and Under-Five-Years Mortality and Nutritional Status in Developing Countries: Evidence from the Demographic and Health Surveys*. DHS Working Papers, 41.

Rutstein, S. O. and Rojas, G. (2003). *Guide to DHS Statistics*. ORC Macro.

Schenker, N. and others. (2002). *Combining estimates from complementary surveys: a case study using prevalence estimates from national health surveys of households and nursing homes*. Public Health Reports, 117, (4), 393.

Schenker, N. and Raghunathan, T. E. (2007). *Combining information from multiple surveys to enhance estimation of measures of health*. Statistics in Medicine, 26, (8), 1802-1811.

Seng, Y. P. (1951). *Hstorical Survey of the Development of Sampling Theories and Practice*. Journal of the Royal Statistical Society, 114, (2), 214-231.

Shaw, S. J. (1978). *The Ottoman Census System and Population, 1831-1914*. International Journal of Middle East Studies, 9, (3), 325-338.

Sneeringer, S. E. (2009). *Fertility Transition in Sub-Saharan Africa: A Comparitive Analysis of Cohort Trends in 30 Countries*. ICF Macro, DHS Comparitive Reports.

Testa, M. R. (2012). *Family Sizes in Europe: Evidence from the 2011 Eurobarometer Survey*. Vienna Institute of Demography of the Austrian Academy of Sciences, European Demographic Research Papers.

Tezcan, S. and Adalı, T. (2012). *Marriage characteristics and reproductive health of adolescents in Turkey: findings from Demographic and Health Surveys 1998 and 2008*. The Turkish Journal of Pediatrics, 54, 273-282.

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

Thomas, S., Sarafin, C. and Simard, M. (2007). *Review of the Weighting Methodology for the Canadian Community Health Survey*. Proceedings of the Federal Committee on Statistical Methodology, Arlington, VA.

Thomas, S. and Wannell, B. (2009). *Combining cycles of the Canadian community health survey*. Statistics Canada Health Reports, 20, (1), 53-58.

Transportation Research Board of the National Academies. (2007). *A Guidebook for Using American Community Survey Data for Transportation Planning*. National Cooperative Highway Research Program.

Tukey, J. W. (1958). *Bias and confidence in not-quite large samples*. The Annals of Mathematical Statistics, 29.

U.S. Census Bureau. (2009). *Design and Methodology, American Community Survey*, U.S. Government Printing Office, Washington, DC.

University of Michigan Population Studies Center and Inter-university Consortuim for Political and Social Research. (2012). *Integrated Fertility Survey Series User Guide*. IFSS Documentation.

Vaessen, M., Thiam, M. and Lê, T. (2005). *Chapter XXII The Demographic and Health Surveys*, United Nations Publications, New York.

Verma, V. (2001). *Leslie Kish: Development of Statistics Internationally*. Statistics Canada Symposium 2001.

Verma, V., Scott, C. and O'Muircheartaigh, C. (1980). *Sample Designs and Sampling Errors for the World Fertility Survey*. Journal of the Royal Statistical Society, 143, (4), 431-473.

Winship, C. and Mare, R. D. (1992). *Models for sample selection bias*. Annual Review of Sociology, 18, 327-350.

Wu, C. (2004). *Combining information from multiple surveys through the empirical likelihood method*. The Canadian Journal of Statistics, 32, (1), 1-12.

Ybarra, L. M. R. and Lohr, S. L. (2008). *Small area estimation when auxiliary information is measured with error*. Biometrika, 95, (4), 919-931.

Zhou, H., Raghunathan, T. E. and Elliott, M. R. (2012). *A Semi-Parametric Approach to Account for Complex Designs in Multiple Imputation*. Federal Committee on Statistical Methodology Research Conference, Washington DC.

Zieschang, K. D. (1990). *Sample Weighting Methods and Estimation of Totals in the Consumer Expenditure Survey*. Journal of the American Statistical Association, 85, (412), 986-1001.

APPENDICES

APPENDIX A: SAS Macro for Mean Number of Children Ever Born

```
%macro mceb(data,pstrata,caseid,byvar,awfdom,survey,intdate);

proc datasets lib=work nolist kill;
quit;
run;

libname tfr 'C:\Users\tugba\Desktop\data preparation\SAS DATA';

/*this step computes the number of CEB until date of interview*/

data work.&data;
set tfr.&data;
/*setting cohorts*/
if (v011 ge 529) and (v011 le 588) then cohort=1;
if (v011 ge 589) and (v011 le 648) then cohort=2;
if (v011 ge 649) and (v011 le 708) then cohort=3;
if (v011 ge 709) and (v011 le 768) then cohort=4;
if (v011 ge 769) and (v011 le 828) then cohort=5;
if (v011 ge 829) and (v011 le 888) then cohort=6;
if (v011 ge 889) and (v011 le 948) then cohort=7;
if (v011 ge 949) and (v011 le 1008) then cohort=8;
if (v011 ge 1009) and (v011 le 1068) then cohort=9;
if (v011 ge 1069) and (v011 le 1128) then cohort=10;
denom=((v005adj3*awfactt)/100);
denomd=((v005adj3*&awfdom)/100);
byear=int(v011/12)+1900;
/*if cohort=&cohort;*/
ceb=0;
%do a=1 %to 9;
if (b3_0&a le &intdate and b3_0&a ne .) then do ceb=ceb+1;
end;
%end;
%do b=10 %to 20;
if (b3_&b le &intdate and b3_&b ne .) then do ceb=ceb+1;
end;
%end;
births=(v005adj3*ceb);
select (TDHS);
when (&survey) output;
otherwise;
end;
run;
```

```
proc sort data=&data;
by &caseid;
run;

DATA &data;
SET &data;
BY &PSTRATA;
TOTAL=1;
RETAIN NPSTRATA 0;
IF FIRST.&PSTRATA THEN NPSTRATA=NPSTRATA+1;
RUN;

/*weights the exposures and births by sample weight.
for exposure, AWF is also included, which is a macro variable for the domain macro*/

proc summary data=&data; /*summarizes the womens dataset - output is called sumwomen*/
VAR births denom denomd;
CLASS tdhs &byvar NPSTRATA SECU cohort ;       /*added byvar*/
OUTPUT OUT=SUMWOMEN SUM=;
RUN;

data CEBNAT;
set SUMWOMEN;
TOTAL=1;
if tdhs=.;
if &byvar=.; /*added byvar*/
if NPSTRATA=.;
if secu=.;
if cohort^=.;
ceb=births/DENOM;
run;

/*this is to get the overall ceb; based on the first surveys age distribution*/

proc summary data=&data;
VAR births denom denomd;
CLASS &byvar tdhs cohort ;
OUTPUT OUT=forad SUM=; /*for weighting the age distribution by the original*/
RUN;

data foradt; /*deleting extra lines*/
set forad;
if &byvar=.;
if tdhs^=.;
if cohort^=.;
run;

data foradt; /*renumbering tdhs, so that we can always use thds1*/
set foradt;
```

```
BY tdhs;
RETAIN tdhsnew 0;
IF FIRST.tdhs THEN tdhsnew=tdhsnew+1;
run;

proc sort data=foradt; /*sorting by cohort to transpose*/
by cohort tdhsnew;
run;

proc transpose data=foradt out=foradt1 prefix=denom; /*transposing to get denom1*/
by &byvar cohort;
id tdhsnew;
var denom;
run;

data cebnat; /*adding denom1 to cebnat, where the cohort cebs are*/
merge cebnat foradt1;
by cohort;
wnum=denom1*ceb;
if denom1^=.;
run;

proc summary data=cebnat; /*summing it up*/
var wnum denom1;
output out=cebnats sum=;
run;

data cebnats; /*computing overall ceb*/
set cebnats;
ceb=wnum/denom1;
total=1;
cohort=100;
run;

data cebnat; /*merging to cebnat as anew line, with cohort=100*/
merge cebnat cebnats;
by cohort;
run;

/*domain version*/

data CEBDOM;
set SUMWOMEN;
if tdhs=.;
if &byvar^=.;
if NPSTRATA=.;
if secu=.;
if cohort^=.;
ceb=births/DENOMD;
```

```
run;

/*this is to get the overall ceb; based on the first surveys age distribution*/

data foradd; /*deleting extra lines*/
set forad;
if &byvar^=.;
if tdhs^=.;
if cohort^=.;
run;

proc sort data=foradd; /*sorting by cohort to transpose*/
by tdhs;
run;

data foradd; /*renumbering tdhs, so that we can always use thds1*/
set foradd;
BY tdhs;
RETAIN tdhsnew 0;
IF FIRST.tdhs THEN tdhsnew=tdhsnew+1;
run;

proc sort data=foradd; /*sorting by cohort to transpose*/
by &byvar cohort tdhsnew;
run;

proc transpose data=foradd out=foradd1 prefix=denomd; /*transposing to get denom1*/
by &byvar cohort;
id tdhsnew;
var denomd;
run;

data cebdom; /*adding denom1 to cebnat, where the cohort cebs are*/
merge cebdom foradd1;
by &byvar cohort;
wnum=denomd1*ceb;
if denomd1^=.;
run;

proc summary data=cebdom; /*summing it up*/
by &byvar;
var wnum denomd1;
output out=cebdoms sum=;
run;

data cebdoms; /*computing overall ceb*/
set cebdoms;
ceb=wnum/denomd1;
cohort=100;
```

```
run;

data cebdom; /*merging to cebnat as anew line, with cohort=100*/
merge cebdom cebdoms;
by &byvar cohort;
run;

data denom;
merge foradt1 foradd1;
by &byvar;
run;


DATA NATIONAL; /*this has strata and cluster level sums*/
 set SUMWOMEN;
IF tdhs=. and &byvar=. and cohort^=. AND NPSTRATA^=. AND SECU^=. THEN OUTPUT
NATIONAL; /*added byvar*/
 RUN;

 DATA DOMAIN; /*ADDED THIS*/
 set SUMWOMEN;
IF tdhs^=. and &byvar^=. and cohort^=. and NPSTRATA^=. AND SECU^=. THEN OUTPUT
DOMAIN;
 RUN;

 DATA FORNOPSU; /*this has strata and cluster level sums*/
 set SUMWOMEN;
IF tdhs=. and &byvar=. and cohort=. and NPSTRATA^=. AND SECU^=. THEN OUTPUT
FORNOPSU; /*added byvar*/
 RUN;

PROC MEANS DATA=FORNOPSU N NOPRINT; /*this creates a strata level data set with a
variable that shows the number of clusters each*/
by NPSTRATA;
VAR secu;
 OUTPUT OUT=NOPSU N=NSECU;
RUN;

 data fornopsu;
 set fornopsu;
 replicate=10*npstrata+secu;
 run;

DATA NATIONAL; /*this step adds the number of clusters information to the cluster level data*/
MERGE NATIONAL nopsu;
BY NPSTRATA;
TOTAL=1;
RUN;
```

```
/*this step creates macro variables from nosecu1 to nosecu(no of strata) */
data _null_;
set nopsu;
suffix=put(_n_,5.);
call symput(cats('nosecu',suffix), NSECU);
run;

/*this step calculates the standard error under stratification and clustering*/

proc sql noprint; /*creating a number of strata variable and calling it nostrata*/
select count(NPSTRATA)
into :nostrata
from work.NOPSU;
quit;

PROC SORT DATA=DOMAIN;
BY NPSTRATA SECU;
RUN;

DATA NATIONAL;
SET NATIONAL;
REPLICATE=NPSTRATA*10+SECU;
RUN;

DATA DOMAIN;
SET DOMAIN;
REPLICATE=NPSTRATA*10+SECU;
RUN;

/*jrr for both national and domains*/

%do i=1 %to &nostrata; /*number of strata-should be defined as a macro variable*/
%do j=1 %to &&nosecu&i; /*number of clusters within strata*/
/*deletes one line*/
data TEMPJRR;
set DOMAIN;
If (NPSTRATA=&i and secu=&j) then delete;
run;

/*doubles the other-same data set*/
data TEMPJRR;
set TEMPJRR;
if (NPSTRATA=&i) then do;
births=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*births;
denom=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*denom;
denomd=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*denomd;
end;
run;
```

```
proc summary data=tempjrr;
VAR denom denomd;
CLASS tdhs &byvar cohort ;
OUTPUT OUT=foradj SUM=; /*for weighting the age distribution by the original*/
RUN;

data foradj; /*deleting extra lines*/
set foradj;
if tdhs^=.;
if cohort^=.;
if &byvar=. then &byvar=200;
run;

proc sort data=foradj; /*sorting by cohort to transpose*/
by tdhs &byvar cohort  ;
run;

data foradj; /*renumbering tdhs, so that we can always use thds1*/
set foradj;
BY tdhs;
RETAIN tdhsnew 0;
IF FIRST.tdhs THEN tdhsnew=tdhsnew+1;
run;

proc sort data=foradj; /*sorting by cohort to transpose*/
by &byvar cohort tdhsnew  ;
run;

proc transpose data=foradj out=foradj1 prefix=denom; /*transposing to get denom1*/
by &byvar cohort;
id tdhsnew ;
var denom;
run;

proc transpose data=foradj out=foradj2 prefix=denomd; /*transposing to get denom1*/
by &byvar cohort;
id tdhsnew ;
var denomd;
run;

data foradj3;
merge foradj1 foradj2;
by &byvar cohort;
run;

PROC SORT DATA=TEMPJRR;
BY &byvar NPSTRATA SECU cohort;
RUN;
```

```
PROC SUMMARY DATA=TEMPJRR;
CLASS &byvar COHORT;/*again same dataset*/
VAR births denom denomd;
OUTPUT OUT=TEMPJRRS SUM=;
run;
DATA TEMPJRRSA; /*This dataset TEMPJRRS is modified to calculate the replicate CEB*/
SET TEMPJRRS;
REPLICATE=&i&j; /*a new variable is created to indicate to which replicate CEB belongs to*/
ceb_jrr=births/denom;
cebd_jrr=births/denomd;
KEEP &byvar ceb_jrr cebd_jrr REPLICATE cohort;
RUN;

data tempjrrsa;
set tempjrrsa;
if cohort^=.;
if &byvar=. then &byvar=200;
run;

proc sort data=tempjrrsa;
by &byvar cohort;
run;

data tempjrrsa1;
merge tempjrrsa foradj3;
by &byvar cohort;
run;

data tempjrrsa1;
set tempjrrsa1;
wnumt=ceb_jrr*denom1;
wnumd=cebd_jrr*denomd1;
run;

proc summary data=tempjrrsa1;
var wnumt wnumd denom1 denomd1;
by &byvar;
output out=tempjrrsa2 sum=;
run;

data tempjrrsa2;
set tempjrrsa2;
cohort=100;
ceb_jrr=wnumt/denom1;
cebd_jrr=wnumd/denomd1;
replicate=&i&j;
drop _type_ _freq_ wnumt wnumd denom1 denomd1;
run;
```

```
data tempjrrsa3;
merge tempjrrsa tempjrrsa2;
by &byvar cohort;
run;

PROC APPEND BASE=DATAJRR DATA=TEMPJRRSA3;  /*The replicate TFRs in
TEMPJRRSs are appended in the DATAJRR dataset*/
RUN;

%end;
%end;

data DATAJRRNAT;
set DATAJRR;
total=1;
if &byvar=200 then output DATAJRRNAT;
drop &byvar cebd_jrr;
run;

data DATAJRRDOM;
set DATAJRR;
if &byvar^=200 then output DATAJRRDOM;
DROP ceb_jrr;
RENAME cebd_jrr=ceb_jrr;
run;

%LET ALLTEMP = NAT DOM;
%LET ALLDATA = NATIONAL DOMAIN;
%LET ALLDOM = TOTAL &BYVAR;

 %DO b=1 %TO  2;

 %LET LOG1=%SCAN(&ALLTEMP,&b);
 %LET LOG2=%SCAN(&ALLDATA,&b);
 %LET LOG3=%SCAN(&ALLDOM,&b);

PROC SORT DATA=DATAJRR&LOG1;
BY &LOG3 cohort;
RUN;

PROC SORT DATA=&LOG2;
BY &LOG3 cohort;
RUN;

data DATAJRR&LOG1; /*other information is brought to the file*/
merge DATAJRR&LOG1 CEB&LOG1;
BY &LOG3 cohort;
npstrata=int(replicate/10);
secu=replicate-10*npstrata;
```

```
KEEP REPLICATE ceb_JRR ceb &LOG3 cohort npstrata secu;
run;

/*data set is ready: randomize according to number of clusters per strata
drop one from all
calculate (tfr_jrr-tfr)^2 and sum - correct jrr*/

/*for domain*/

PROC SORT DATA=DATAJRRDOM;
BY NPSTRATA SECU;
RUN;
data _null_; /*this creates as many random numbers as the number of strata*/
set nopsu;
RN=int(1+NSECU*ranuni(3674));
suffix=put(_n_,5.);
call symput(cats('RN',suffix), RN);
run;

/*the DATAJRR file is reduced*/
data DATAJRR2&LOG1;
set DATAJRR&LOG1;
%do i=1 %to &nostrata;
if replicate=(%eval(&i)*10+%eval(&&RN&i)) /*(NPSTRATA=&i and secu=&&RN&i)*/ then
delete;
%end;
run;

data DATAJRR2&LOG1;
set DATAJRR2&LOG1;
dif2=(ceb-ceb_JRR)**2;
run;

PROC SORT DATA=DATAJRR2&LOG1;
BY &LOG3 cohort;
RUN;

proc summary data=DATAJRR2&LOG1;
CLASS &LOG3 cohort;
var dif2;
output out=CJRR&LOG1 sum=cebcvar;
run;

data CJRR&LOG1; /*final complex sample standard error file*/
set CJRR&LOG1;
IF &b=1 and total=. or cohort=. then delete;
if &b=2 and &byvar=. or cohort=. then delete;
cebcSE=(cebcvar)**0.5;
drop _type_ _freq_;
```

```
run;

%end;

/*prep for SRS*/

DATA &data;
SET &data;
BY &CASEID;
TOTAL=1;
RETAIN COUNT 0;
IF FIRST.&CASEID THEN COUNT=COUNT+1;
RUN;

DATA &data;
SET &data;
BY v001new;
RETAIN NCLUSTER 0;
IF FIRST.v001new THEN NCLUSTER=NCLUSTER+1;
RUN;

proc sql noprint; /*creating the maximum number of women variable*/
select max(count), min(count), max(ncluster)
into: stop, :start, :ncluster
from work.&data;
quit;

/*SRS FOR NATIONAL*/

 %DO j=&start %TO &stop; /*This goes from 1 to number of women*/ /*&START;  /*SO THAT
THE LOOP WILL WORK ONLY ONCE*/
      DATA TEMPW;
       SET &data;
       IF COUNT=&j THEN DELETE;  /* deletes one woman at a time */
      RUN;

PROC SUMMARY DATA=TEMPW;
CLASS tdhs &BYVAR cohort;/*again same dataset*/
VAR births denom denomd;
OUTPUT OUT=TEMPW2 SUM=;
RUN;

data tempw2h;
set tempw2;
if tdhs^=.;
if cohort^=.;
if &byvar=. then &byvar=200;
run;
```

```
proc sort data=tempw2h;
by tdhs &byvar cohort;
run;

data tempw2h;
set tempw2h;
by tdhs;
retain tdhsnew 0;
if first.tdhs then tdhsnew=tdhsnew+1;
run;

proc sort data=tempw2h;
by &byvar cohort tdhsnew;
run;

proc transpose data=tempw2h out=tempw2h2 prefix=denom;
by &byvar cohort;
id tdhsnew;
var denom;
run;

proc transpose data=tempw2h out=tempw2h3 prefix=denomd;
by &byvar cohort;
id tdhsnew;
var denomd;
run;

data tempw2h4;
merge tempw2h2 tempw2h3;
by &byvar cohort;
run;

proc sort data=tempw;
by &byvar cohort;
run;

proc summary data=tempw;
class &byvar cohort;
var births denom denomd;
output out=tempw2g sum=;
run;

DATA TEMPW2i; /*This dataset TEMPJRRS is modified to calculate the replicate TFR*/
SET TEMPW2g;
TOTAL=1;
REPLICATE=&j; /*a new variable is created to indicate to which replicate TFR belongs to*/
ceb_srs=births/denom;
cebd_srs=births/denomd;
KEEP ceb_srs cebd_srs REPLICATE TOTAL &BYVAR cohort;
```

```
data tempw2i;
set tempw2i;
if cohort^=.;
if &byvar=. then &byvar=200;

proc sort data=tempw2i;
by &byvar cohort;
run;

data tempw2j;
merge tempw2i tempw2h4;
by &byvar cohort;
run;

data tempw2j;
set tempw2j;
wnumt=ceb_srs*denom1;
wnumd=cebd_srs*denomd1;
run;

proc summary data=tempw2j;
var wnumt wnumd denom1 denomd1;
by &byvar;
output out=tempw2k sum=;
run;

data tempw2k;
set tempw2k;
cohort=100;
ceb_srs=wnumt/denom1;
cebd_srs=wnumt/denomd1;
replicate=&j;
drop _type_ _freq_ wnumt wnumd denom1 denomd1;
run;

data tempw2s;
merge tempw2i tempw2k;
by &byvar cohort;
run;

PROC APPEND BASE=DATAWTOT DATA=TEMPW2s;   /*The replicate TFRs in
TEMPJRRSs are appended in the DATAJRR dataset*/
RUN;

%end;

data _null_; /*this creates as many random numbers as the number of cases*/
set &data;
```

```
RN=int(1+count*ranuni(1));
call symput('RN', RN);
run;

/*the DATAW file is reduced*/
data DATAWTOT;
set DATAWTOT;
if (replicate=&RN) then delete;
run;

DATA DATAW;
SET DATAWTOT;
IF &BYVAR=200;
total=1;
DROP cebd_srs ;
RUN;

DATA DATAWD;
SET DATAWTOT ;
IF &BYVAR^=200;
DROP ceb_srs ;
RENAME cebd_srs=ceb_srs ;
RUN;

%LET ALLDATA2 = DATAW DATAWD;

 %DO i=1 %TO  2;

 %LET LOG1=%SCAN(&ALLTEMP,&i);
 %LET LOG2=%SCAN(&ALLDATA2,&i);
 %LET LOG3=%SCAN(&ALLDOM,&i);

PROC SORT DATA=&LOG2;
BY &LOG3 cohort;
RUN;

DATA &LOG2;
MERGE &LOG2 CEB&LOG1 ;
BY &LOG3 cohort;
srsdif2=(ceb-ceb_srs)**2;
RUN;

 PROC SUMMARY DATA=&LOG2 SUM;
  CLASS &LOG3 cohort;
  VAR srsdif2;
  OUTPUT OUT=RESULT&LOG1 SUM=cebsrsvar;
 RUN;

DATA RESULT&LOG1;
```

```
SET RESULT&LOG1;
  IF &LOG3=. or cohort=. THEN DELETE;
  cebsrsse=(cebsrsvar)**0.5;
  drop _type_ _freq_;
 RUN;


%IF &BYVAR^= %THEN %DO; /*&BYVAR replaced by v024*/
                      %LET dsid=%SYSFUNC(OPEN(&DATA,i));
                      %LET n   =%SYSFUNC(VARNUM(&dsid,&BYVAR));
                      %LET fmt =%SYSFUNC(VARFMT(&dsid,&n));
                      %PUT &fmt;
                      %LET rc=%SYSFUNC(CLOSE(&dsid));
                   %END;


DATA FINAL&LOG1;
MERGE RESULT&LOG1 CJRR&LOG1 ceb&LOG1;
 BY &LOG3;

  IF cebSRSSE^=0 THEN DEFT = cebcse/(cebsrsse);
ELSE DEFT=.;

IF ceb^=0 THEN RELERROR = cebCSE/ceb;
ELSE RELERROR=.;

 LOWER =ceb -2*cebCSE;  UPPER= ceb +2*cebcSE;
 IF LOWER<0 THEN LOWER=0;

 LABEL='mean number of children ever born';

  TYPE='Mean';

 %IF &BYVAR^=   %THEN %DO;
FORMAT &LOG3 &fmt.;
%END;
CLUSTER=&NCLUSTER;
STRATA=&NOSTRATA;
WOMEN=&stop;

KEEP LABEL TYPE ceb cebCSE cebSRSSE DEFT RELERROR LOWER UPPER &LOG3
CLUSTER STRATA WOMEN;
RUN;


PROC FORMAT;
VALUE NAF 999999999='NA' 0='Entire sample';
RUN;
```

```
OPTIONS LS=120 NONUMBER NODATE FORMDLIM=' ';

PROC SORT DATA=FINAL&LOG1 NODUP;
BY &LOG3;
RUN;

DATA FINAL&LOG1;
SET FINAL&LOG1;
IF &i=1 THEN &LOG3=0;
IF &i=2 AND &LOG3=. THEN DELETE;
RENAME &LOG3=LEVEL;
RUN;

PROC APPEND BASE=TABLE DATA=FINAL&LOG1;  /*The replicate TFRs in TEMPJRRSs
are appended in the DATAJRR dataset*/
RUN;

%END;

ods html;

PROC PRINT DATA=TABLE WIDTH=MINIMUM HEADING=HORIZONTAL NOOBS;
 VAR LABEL TYPE ceb cebCSE LOWER UPPER cebSRSSE DEFT RELERROR
 CLUSTER STRATA WOMEN;
 BY LEVEL;
  FORMAT LEVEL NAF.;
TITLE "&byvar dhs=&survey";
RUN;

ods html close;

%mend;
```

APPENDIX B: SAS Macro for Total Fertility Rate

```
%macro nationaltfr(data,intdate,months,pstrata,caseid,byvar,inflate,survey);

proc datasets lib=work nolist kill;
quit;
run;

libname tfr 'C:\Users\tugba\Desktop\data preparation\SAS DATA';

data &data;
set tfr.&data;
do  BIR1519=0;  BIR2024=0;  BIR2529=0;  BIR3034=0;  BIR3539=0;  BIR4044=0;  BIR4549=0;
/*initilalize birth variables*/
end;
intdate=&intdate;
```

```
select (TDHS);
when (&survey) output;
otherwise;
end;
run;

proc sort data=&data;
by &caseid;
run;

DATA &data;
SET &data;
BY &PSTRATA;
TOTAL=1;
RETAIN NPSTRATA 0;
IF FIRST.&PSTRATA THEN NPSTRATA=NPSTRATA+1;
RUN;


proc sql noprint; /*creating the maximum number of children variable*/
select max(v201)
into :max
from work.&data;
quit;

%macro birth1; /*for births from 1 to 9*/
%let ini=180;
data &data;
set &data;

%do a=15 %to 45 %by 5;

if ((b3_0&i ge (intdate-%eval(&months))) and (b3_0&i<intdate) and ((b3_0&i-v011) ge
%eval(&ini)) and ((b3_0&i-v011)<%eval(&ini+60)))
then do BIR&a%eval(&a+4)=BIR&a%eval(&a+4)+1;
end;

else do BIR&a%eval(&a+4)=BIR&a%eval(&a+4);
end;

%let ini=%eval(&ini+60);
%end;
run;
%mend;

%macro birth2; /*for births from 10 to max or 20*/
%let ini=180;
data &data;
set &data;
```

```
%do a=15 %to 45 %by 5;

if ((b3_&j ge (intdate-%eval(&months))) and (b3_&j<intdate) and ((b3_&j-v011) ge %eval(&ini))
and ((b3_&j-v011)<%eval(&ini+60)))
then do BIR&a%eval(&a+4)=BIR&a%eval(&a+4)+1;
end;

else do BIR&a%eval(&a+4)=BIR&a%eval(&a+4);
end;

%let ini=%eval(&ini+60);
%end;
run;
%mend;

%macro addbirth; /*adds 7 variables by running the macros above: the number of births in the past
3 years by age*/
%if (&max<=9) %then %let stop1=&max;
%else %do; %let stop1=9; %let stop2=&max; %end;

%do i=1 %to &stop1;
%birth1
%end;

%if (&max>9) %then %do j=10 %to &stop2;
%birth2
%end;

%mend;

%addbirth;

/*options mlogic symbolgen mprint;*/

/*the exposures are added to the women's data set*/
DATA &DATA;
 SET &DATA; /*AWFACTT ADDED*/

 AGE_MAX = intdate-V011;            /* mom's age at interview */
 AGE_MIN = intdate-V011-&months;        /* mom's age at start of period */

  %DO i=15 %TO 45 %BY 5;        /*  15 20 25 30 35 40 45 */
 %LET j=%EVAL(&i+4);          /*  19 24 29 34 39 44 49 */
 %LET LOW=%EVAL(12*&i);
 %LET HIGH=%EVAL(12*(&i+5) - 1);
 %LET STOP=%EVAL(&HIGH+1);

 EXP_&i&j=0;
```

```
 IF AGE_MAX >&HIGH AND AGE_MIN<&STOP
          THEN EXP_&i&j= (&STOP - MAX(&LOW,AGE_MIN))/12;

                              IF &LOW<=AGE_MAX<=&HIGH
          THEN EXP_&i&j= (AGE_MAX - MAX(&LOW, AGE_MIN))/12;

 %END;

RUN;

/*weights the exposures and births by sample weight.
for exposure, AWF is also included, which is a macro variable for the domain macro*/
data &data;
set &data;
%do i=15 %to 45 %by 5;
W_BIR&i%eval(&i+4)=BIR&i%eval(&i+4)*v005adj3;                  /*changed both weights to new
ones!*/
W_EXP&i%eval(&i+4)=EXP_&i%eval(&i+4)*v005adj3*(awfactt/100);
W_EXPD&i%eval(&i+4)=EXP_&i%eval(&i+4)*v005adj3*(&inflate/100);
%end;
run;

proc summary data=&data; /*summarizes the womens dataset - output is called sumwomen*/
VAR W_BIR1519 W_BIR2024 W_BIR2529 W_BIR3034 W_BIR3539 W_BIR4044
W_BIR4549 W_EXP1519 W_EXP2024 W_EXP2529 W_EXP3034 W_EXP3539 W_EXP4044
W_EXP4549
W_EXPD1519 W_EXPD2024 W_EXPD2529 W_EXPD3034 W_EXPD3539 W_EXPD4044
W_EXPD4549;
CLASS &byvar NPSTRATA SECU;            /*added byvar*/
OUTPUT OUT=SUMWOMEN SUM=;
RUN;

data TFRNAT; /*national tfr is in this file: NATTFR*/
set SUMWOMEN;
TOTAL=1;
if &byvar=.; /*added byvar*/
if NPSTRATA=.;
if secu=.;
IF W_EXP1519^=0 THEN ASFR1519= W_BIR1519/W_EXP1519; ELSE ASFR1519=.;
 IF W_EXP2024^=0 THEN ASFR2024= W_BIR2024/W_EXP2024; ELSE ASFR2024=.;
 IF W_EXP2529^=0 THEN ASFR2529= W_BIR2529/W_EXP2529; ELSE ASFR2529=.;
 IF W_EXP3034^=0 THEN ASFR3034= W_BIR3034/W_EXP3034; ELSE ASFR3034=.;
 IF W_EXP3539^=0 THEN ASFR3539= W_BIR3539/W_EXP3539; ELSE ASFR3539=.;
 IF W_EXP4044^=0 THEN ASFR4044= W_BIR4044/W_EXP4044; ELSE ASFR4044=.;
 IF W_EXP4549^=0 THEN ASFR4549= W_BIR4549/W_EXP4549; ELSE ASFR4549=.;
 TFR=SUM(ASFR1519,ASFR2024,ASFR2529,ASFR3034,
        ASFR3539,ASFR4044,ASFR4549)*5;
W_N=SUM(W_EXP1519,W_EXP2024,W_EXP2529,W_EXP3034,W_EXP3539,W_EXP4044,
W_EXP4549);
```

```
 run;

data TFRDOM; /*added this*/
set SUMWOMEN;
if &byvar^=.; /*added byvar*/
if NPSTRATA=.;
if secu=.;
IF W_EXPD1519^=0 THEN ASFR1519= W_BIR1519/W_EXPD1519; ELSE ASFR1519=.;
 IF W_EXPD2024^=0 THEN ASFR2024= W_BIR2024/W_EXPD2024; ELSE ASFR2024=.;
 IF W_EXPD2529^=0 THEN ASFR2529= W_BIR2529/W_EXPD2529; ELSE ASFR2529=.;
 IF W_EXPD3034^=0 THEN ASFR3034= W_BIR3034/W_EXPD3034; ELSE ASFR3034=.;
 IF W_EXPD3539^=0 THEN ASFR3539= W_BIR3539/W_EXPD3539; ELSE ASFR3539=.;
 IF W_EXPD4044^=0 THEN ASFR4044= W_BIR4044/W_EXPD4044; ELSE ASFR4044=.;
 IF W_EXPD4549^=0 THEN ASFR4549= W_BIR4549/W_EXPD4549; ELSE ASFR4549=.;
 TFR=SUM(ASFR1519,ASFR2024,ASFR2529,ASFR3034,
        ASFR3539,ASFR4044,ASFR4549)*5;
W_N=SUM(W_EXP1519,W_EXP2024,W_EXP2529,W_EXP3034,W_EXP3539,W_EXP4044,
W_EXP4549);
                          run;

DATA NATIONAL; /*this has strata and cluster level sums*/
 set SUMWOMEN;
IF &byvar=. and NPSTRATA^=. AND SECU^=. THEN OUTPUT NATIONAL; /*added byvar*/
 RUN;

 DATA DOMAIN;
 set SUMWOMEN;
IF &byvar^=. and NPSTRATA^=. AND SECU^=. THEN OUTPUT DOMAIN;
 RUN;

PROC MEANS DATA=NATIONAL N NOPRINT; /*this creates a strata level data set with a
variable that shows the number of clusters each*/
by NPSTRATA;
VAR secu;
 OUTPUT OUT=NOPSU N=NSECU;
RUN;

DATA NATIONAL; /*this step adds the number of clusters information to the cluster level data*/
MERGE NATIONAL nopsu;
BY NPSTRATA;
TOTAL=1;
RUN;

/*this step creates macro variables from nosecu1 to nosecu(no of strata) */
data _null_;
set nopsu;
suffix=put(_n_,5.);
call symput(cats('nosecu',suffix), NSECU);
run;
```

```
/*this step calculates the standard error under stratification and clustering*/

proc sql noprint; /*creating a number of strata variable and calling it nostrata*/
select count(NPSTRATA)
into :nostrata
from work.NOPSU;
quit;

PROC SORT DATA=DOMAIN;
BY NPSTRATA SECU;
RUN;

DATA NATIONAL;
SET NATIONAL;
REPLICATE=NPSTRATA*10+SECU;
RUN;

DATA DOMAIN;
SET DOMAIN;
REPLICATE=NPSTRATA*10+SECU;
RUN;

/*jrr for both national and domains*/

%do i=1 %to &nostrata; /*number of strata-should be defined as a macro variable*/
%do j=1 %to &&nosecu&i; /*number of clusters within strata*/

/*deletes one line*/
data TEMPJRR;
set DOMAIN;
If (NPSTRATA=&i and secu=&j) then delete;
run;

/*doubles the other-same data set*/
data TEMPJRR;
set TEMPJRR;
if (NPSTRATA=&i) then do;
W_BIR1519=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_BIR1519;
W_BIR2024=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_BIR2024;
W_BIR2529=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_BIR2529;
W_BIR3034=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_BIR3034;
W_BIR3539=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_BIR3539;
W_BIR4044=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_BIR4044;
W_BIR4549=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_BIR4549;
W_EXP1519=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_EXP1519;
W_EXP2024=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_EXP2024;
W_EXP2529=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_EXP2529;
W_EXP3034=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_EXP3034;
W_EXP3539=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_EXP3539;
```

```
W_EXP4044=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_EXP4044;
W_EXP4549=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_EXP4549;
W_EXPD1519=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_EXPD1519;
W_EXPD2024=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_EXPD2024;
W_EXPD2529=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_EXPD2529;
W_EXPD3034=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_EXPD3034;
W_EXPD3539=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_EXPD3539;
W_EXPD4044=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_EXPD4044;
W_EXPD4549=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*W_EXPD4549;

end;
run;

PROC SORT DATA=TEMPJRR;
BY &BYVAR NPSTRATA SECU;
RUN;

PROC SUMMARY DATA=TEMPJRR;
CLASS &BYVAR;/*again same dataset*/
VAR W_BIR1519 W_BIR2024 W_BIR2529 W_BIR3034 W_BIR3539 W_BIR4044
W_BIR4549 W_EXP1519 W_EXP2024 W_EXP2529 W_EXP3034 W_EXP3539 W_EXP4044
W_EXP4549
W_EXPD1519 W_EXPD2024 W_EXPD2529 W_EXPD3034 W_EXPD3539 W_EXPD4044
W_EXPD4549;
OUTPUT OUT=TEMPJRRS SUM=;

DATA TEMPJRRS; /*This dataset TEMPJRRS is modified to calculate the replicate TFR*/
SET TEMPJRRS;
REPLICATE=&i&j; /*a new variable is created to indicate to which replicate TFR belongs to*/
IF W_EXP1519^=0 THEN ASFR1C= W_BIR1519/W_EXP1519; ELSE ASFR1C=.;
IF W_EXP2024^=0 THEN ASFR2C= W_BIR2024/W_EXP2024; ELSE ASFR2C=.;
IF W_EXP2529^=0 THEN ASFR3C= W_BIR2529/W_EXP2529; ELSE ASFR3C=.;
IF W_EXP3034^=0 THEN ASFR4C= W_BIR3034/W_EXP3034; ELSE ASFR4C=.;
IF W_EXP3539^=0 THEN ASFR5C= W_BIR3539/W_EXP3539; ELSE ASFR5C=.;
IF W_EXP4044^=0 THEN ASFR6C= W_BIR4044/W_EXP4044; ELSE ASFR6C=.;
IF W_EXP4549^=0 THEN ASFR7C= W_BIR4549/W_EXP4549; ELSE ASFR7C=.;
IF W_EXPD1519^=0 THEN ASFR1CD= W_BIR1519/W_EXPD1519; ELSE ASFR1CD=.;
IF W_EXPD2024^=0 THEN ASFR2CD= W_BIR2024/W_EXPD2024; ELSE ASFR2CD=.;
IF W_EXPD2529^=0 THEN ASFR3CD= W_BIR2529/W_EXPD2529; ELSE ASFR3CD=.;
IF W_EXPD3034^=0 THEN ASFR4CD= W_BIR3034/W_EXPD3034; ELSE ASFR4CD=.;
IF W_EXPD3539^=0 THEN ASFR5CD= W_BIR3539/W_EXPD3539; ELSE ASFR5CD=.;
IF W_EXPD4044^=0 THEN ASFR6CD= W_BIR4044/W_EXPD4044; ELSE ASFR6CD=.;
IF W_EXPD4549^=0 THEN ASFR7CD= W_BIR4549/W_EXPD4549; ELSE ASFR7CD=.;
TFR_JRR=SUM(ASFR1C,ASFR2C,ASFR3C,ASFR4C,ASFR5C,ASFR6C,ASFR7C)*5;
TFR_JRRD=SUM(ASFR1CD,ASFR2CD,ASFR3CD,ASFR4CD,ASFR5CD,ASFR6CD,ASFR7C
D)*5;
                         KEEP &BYVAR ASFR1C ASFR2C ASFR3C ASFR4C ASFR5C
ASFR6C ASFR7C
ASFR1CD ASFR2CD ASFR3CD ASFR4CD ASFR5CD ASFR6CD ASFR7CD
```

```
              TFR_JRR TFR_JRRD REPLICATE;
RUN;
PROC APPEND BASE=DATAJRR DATA=TEMPJRRS;  /*The replicate TFRs in TEMPJRRSs
are appended in the DATAJRR dataset*/
RUN;

%end;
%end;


data DATAJRRNAT;
set DATAJRR;
total=1;
if &byvar=. then output DATAJRRNAT;
drop &byvar ASFR1CD ASFR2CD ASFR3CD ASFR4CD ASFR5CD ASFR6CD ASFR7CD
TFR_JRRD;
run;

data DATAJRRDOM;
set DATAJRR;
if &byvar^=. then output DATAJRRDOM;
DROP ASFR1C ASFR2C ASFR3C ASFR4C ASFR5C ASFR6C ASFR7C TFR_JRR;
RENAME      ASFR1CD=ASFR1C      ASFR2CD=ASFR2C      ASFR3CD=ASFR3C
ASFR4CD=ASFR4C    ASFR5CD=ASFR5C    ASFR6CD=ASFR6C    ASFR7CD=ASFR7C
TFR_JRRD=TFR_JRR;
run;

PROC SORT DATA=DATAJRRDOM;
BY &BYVAR;
RUN;

PROC SORT DATA=DOMAIN;
BY &BYVAR;
RUN;

%LET ALLTEMP = NAT DOM;
%LET ALLDATA = NATIONAL DOMAIN;
%LET ALLDOM = TOTAL &BYVAR;

 %DO b=1 %TO  2;

 %LET LOG1=%SCAN(&ALLTEMP,&b);
 %LET LOG2=%SCAN(&ALLDATA,&b);
 %LET LOG3=%SCAN(&ALLDOM,&b);

data DATAJRR&LOG1; /*other information is brought to the file*/
merge DATAJRR&LOG1 TFR&LOG1;
BY &LOG3;
KEEP ASFR1519 ASFR2024 ASFR2529 ASFR3034 ASFR3539 ASFR4044 ASFR4549
```

```
  ASFR1C ASFR2C ASFR3C ASFR4C ASFR5C ASFR6C ASFR7C
  REPLICATE TFR_JRR TFR &LOG3 W_N;
run;

PROC SORT DATA=DATAJRR&LOG1;
BY REPLICATE;
RUN;

PROC SORT DATA=&LOG2;
BY REPLICATE;
RUN;

data DATAJRR&LOG1; /*other information is brought to the file*/
merge DATAJRR&LOG1 &LOG2;
BY REPLICATE;
KEEP ASFR1519 ASFR2024 ASFR2529 ASFR3034 ASFR3539 ASFR4044 ASFR4549
  ASFR1C ASFR2C ASFR3C ASFR4C ASFR5C ASFR6C ASFR7C
  REPLICATE TFR_JRR TFR NPSTRATA SECU &LOG3 W_N;
run;

/*data set is ready: randomize according to number of clusters per strata
drop one from all
calculate (tfr_jrr-tfr)^2 and sum - correct jrr*/

/*for domains*/

PROC SORT DATA=DATAJRRDOM;
BY NPSTRATA SECU;
RUN;

data _null_; /*this creates as many random numbers as the number of strata*/
set nopsu;
RN=int(1+NSECU*ranuni(3674));
suffix=put(_n_,5.);
call symput(cats('RN',suffix), RN);
run;

/*the DATAJRR file is reduced*/
data DATAJRR2&LOG1;
set DATAJRR&LOG1;
%do i=1 %to &nostrata;
if (NPSTRATA=&i and secu=&&RN&i) then delete;
%end;

data DATAJRR2&LOG1;
set DATAJRR2&LOG1;
TFRCD=(TFR-TFR_JRR)**2;
ASFR1CD=(ASFR1519-ASFR1C)**2;
ASFR2CD=(ASFR2024-ASFR2C)**2;
```

```
ASFR3CD=(ASFR2529-ASFR3C)**2;
ASFR4CD=(ASFR3034-ASFR4C)**2;
ASFR5CD=(ASFR3539-ASFR5C)**2;
ASFR6CD=(ASFR4044-ASFR6C)**2;
ASFR7CD=(ASFR4549-ASFR7C)**2;
run;

PROC SORT DATA=DATAJRR2&LOG1;
BY &LOG3;
RUN;

proc summary data=DATAJRR2&LOG1;
CLASS &LOG3;
var TFRCD ASFR1CD ASFR2CD ASFR3CD ASFR4CD ASFR5CD ASFR6CD ASFR7CD;
output out=CJRR&LOG1 sum=TFRCVAR ASFR1CVAR ASFR2CVAR ASFR3CVAR
ASFR4CVAR ASFR5CVAR ASFR6CVAR ASFR7CVAR;
run;
data CJRR&LOG1; /*final complex sample standard error file*/
set CJRR&LOG1;
IF &b=1 and total=. then delete;
if &b=2 and &byvar=. then delete;
TFRCSE=(TFRCVAR)**0.5;
ASFR1CSE=(ASFR1CVAR)**0.5;
ASFR2CSE=(ASFR2CVAR)**0.5;
ASFR3CSE=(ASFR3CVAR)**0.5;
ASFR4CSE=(ASFR4CVAR)**0.5;
ASFR5CSE=(ASFR5CVAR)**0.5;
ASFR6CSE=(ASFR6CVAR)**0.5;
ASFR7CSE=(ASFR7CVAR)**0.5;
drop _type_ _freq_;
run;

%end;

/*prep for SRS*/

DATA &data;
SET &data;
BY &CASEID;
TOTAL=1;
RETAIN COUNT 0;
IF FIRST.&CASEID THEN COUNT=COUNT+1;
RUN;

DATA &data;
SET &data;
BY v001new;
RETAIN NCLUSTER 0;
IF FIRST.v001new THEN NCLUSTER=NCLUSTER+1;
```

```
RUN;

proc sql noprint; /*creating the maximum number of women variable*/
select max(count), min(count), max(ncluster)
into: stop, :start, :ncluster
from work.&data;
quit;

/*SRS FOR NATIONAL*/

 %DO j=&start %TO &stop; /*This goes from 1 to number of women*/
      DATA TEMPW;
       SET &data;
       IF COUNT=&j THEN DELETE;  /* deletes one woman at a time */
      RUN;

PROC SUMMARY DATA=TEMPW;
CLASS &BYVAR;/*again same dataset*/
VAR W_BIR1519 W_BIR2024 W_BIR2529 W_BIR3034 W_BIR3539 W_BIR4044
W_BIR4549 W_EXP1519 W_EXP2024 W_EXP2529 W_EXP3034 W_EXP3539 W_EXP4044
W_EXP4549
W_EXPD1519 W_EXPD2024 W_EXPD2529 W_EXPD3034 W_EXPD3539 W_EXPD4044
W_EXPD4549;
OUTPUT OUT=TEMPW2 SUM=;

DATA TEMPW2; /*This dataset TEMPJRRS is modified to calculate the replicate TFR*/
SET TEMPW2;
TOTAL=1;
REPLICATE=&j; /*a new variable is created to indicate to which replicate TFR belongs to*/
IF W_EXP1519^=0 THEN ASFR1SRS= W_BIR1519/W_EXP1519; ELSE ASFR1SRS=.;
IF W_EXP2024^=0 THEN ASFR2SRS= W_BIR2024/W_EXP2024; ELSE ASFR2SRS=.;
IF W_EXP2529^=0 THEN ASFR3SRS= W_BIR2529/W_EXP2529; ELSE ASFR3SRS=.;
IF W_EXP3034^=0 THEN ASFR4SRS= W_BIR3034/W_EXP3034; ELSE ASFR4SRS=.;
IF W_EXP3539^=0 THEN ASFR5SRS= W_BIR3539/W_EXP3539; ELSE ASFR5SRS=.;
IF W_EXP4044^=0 THEN ASFR6SRS= W_BIR4044/W_EXP4044; ELSE ASFR6SRS=.;
IF W_EXP4549^=0 THEN ASFR7SRS= W_BIR4549/W_EXP4549; ELSE ASFR7SRS=.;
IF W_EXPD1519^=0 THEN ASFR1SRSD= W_BIR1519/W_EXPD1519; ELSE ASFR1SRSD=.;
IF W_EXPD2024^=0 THEN ASFR2SRSD= W_BIR2024/W_EXPD2024; ELSE ASFR2SRSD=.;
IF W_EXPD2529^=0 THEN ASFR3SRSD= W_BIR2529/W_EXPD2529; ELSE ASFR3SRSD=.;
IF W_EXPD3034^=0 THEN ASFR4SRSD= W_BIR3034/W_EXPD3034; ELSE ASFR4SRSD=.;
IF W_EXPD3539^=0 THEN ASFR5SRSD= W_BIR3539/W_EXPD3539; ELSE ASFR5SRSD=.;
IF W_EXPD4044^=0 THEN ASFR6SRSD= W_BIR4044/W_EXPD4044; ELSE ASFR6SRSD=.;
IF W_EXPD4549^=0 THEN ASFR7SRSD= W_BIR4549/W_EXPD4549; ELSE ASFR7SRSD=.;
TFR_PARTW=SUM(ASFR1SRS,ASFR2SRS,ASFR3SRS,ASFR4SRS,
              ASFR5SRS,ASFR6SRS,ASFR7SRS)*5;
TFR_PARTWD=SUM(ASFR1SRSD,ASFR2SRSD,ASFR3SRSD,ASFR4SRSD,
              ASFR5SRSD,ASFR6SRSD,ASFR7SRSD)*5;
                                    KEEP TFR_PARTW ASFR1SRS ASFR2SRS ASFR3SRS
ASFR4SRS
```

```
        ASFR5SRS ASFR6SRS ASFR7SRS TFR_PARTWD ASFR1SRSD ASFR2SRSD
ASFR3SRSD ASFR4SRSD
        ASFR5SRSD ASFR6SRSD ASFR7SRSD REPLICATE TOTAL &BYVAR;
RUN;
PROC APPEND BASE=DATAWTOT DATA=TEMPW2;  /*The replicate TFRs in TEMPJRRSs
are appended in the DATAJRR dataset*/
RUN;

%end;

data _null_; /*this creates as many random numbers as the number of strata*/
set &data;
RN=int(1+count*ranuni(1));
call symput('RN', RN);
run;

/*the DATAW file is reduced*/
data DATAWTOT;
set DATAWTOT;
if (replicate=&RN) then delete;
run;

DATA DATAW;
SET DATAWTOT;
DROP TFR_PARTWD ASFR1SRSD ASFR2SRSD ASFR3SRSD ASFR4SRSD
        ASFR5SRSD ASFR6SRSD ASFR7SRSD ;
IF &BYVAR=.;
RUN;

DATA DATAWD;
SET DATAWTOT ;
DROP TFR_PARTW ASFR1SRS ASFR2SRS ASFR3SRS ASFR4SRS
        ASFR5SRS ASFR6SRS ASFR7SRS;
RENAME TFR_PARTWD=TFR_PARTW ASFR1SRSD=ASFR1SRS
ASFR2SRSD=ASFR2SRS ASFR3SRSD=ASFR3SRS ASFR4SRSD=ASFR4SRS
        ASFR5SRSD=ASFR5SRS ASFR6SRSD=ASFR6SRS ASFR7SRSD=ASFR7SRS ;
IF &BYVAR^=.;
RUN;

PROC SORT DATA=DATAWD;
BY &BYVAR;
RUN;

%LET ALLDATA2 = DATAW DATAWD;

 %DO i=1 %TO  2;

 %LET LOG1=%SCAN(&ALLTEMP,&i);
 %LET LOG2=%SCAN(&ALLDATA2,&i);
```

```
 %LET LOG3=%SCAN(&ALLDOM,&i);

DATA &LOG2;
MERGE &LOG2 TFR&LOG1 ;
BY &LOG3;
TFRSRSD=(TFR-TFR_PARTW)**2;
ASFR1SRSD=(ASFR1519-ASFR1SRS)**2;
ASFR2SRSD=(ASFR2024-ASFR2SRS)**2;
ASFR3SRSD=(ASFR2529-ASFR3SRS)**2;
ASFR4SRSD=(ASFR3034-ASFR4SRS)**2;
ASFR5SRSD=(ASFR3539-ASFR5SRS)**2;
ASFR6SRSD=(ASFR4044-ASFR6SRS)**2;
ASFR7SRSD=(ASFR4549-ASFR7SRS)**2;
RUN;


 PROC SUMMARY DATA=&LOG2 SUM;
  CLASS &LOG3;
  VAR TFRSRSD ASFR1SRSD ASFR2SRSD ASFR3SRSD ASFR4SRSD ASFR5SRSD
ASFR6SRSD ASFR7SRSD;
  OUTPUT OUT=RESULT&LOG1 SUM=TFRSRSV ASFR1SRSV ASFR2SRSV ASFR3SRSV
ASFR4SRSV ASFR5SRSV ASFR6SRSV ASFR7SRSV;
 RUN;

DATA RESULT&LOG1;
SET RESULT&LOG1;
  IF &LOG3=. THEN DELETE;
  TFRSRSSE=(TFRSRSV)**0.5;
  ASFR1SRSSE=(ASFR1SRSV)**0.5;
  ASFR2SRSSE=(ASFR2SRSV)**0.5;
  ASFR3SRSSE=(ASFR3SRSV)**0.5;
  ASFR4SRSSE=(ASFR4SRSV)**0.5;
  ASFR5SRSSE=(ASFR5SRSV)**0.5;
  ASFR6SRSSE=(ASFR6SRSV)**0.5;
  ASFR7SRSSE=(ASFR7SRSV)**0.5;
  KEEP TFRSRSSE ASFR1SRSSE ASFR2SRSSE ASFR3SRSSE ASFR4SRSSE ASFR5SRSSE
ASFR6SRSSE ASFR7SRSSE &LOG3;
 RUN;

PROC SORT DATA=RESULT&LOG1;
BY &LOG3;  /*&LOG3 replaced by TOTAL*/
RUN;

%IF &BYVAR^= %THEN %DO; /*&BYVAR replaced by v024*/
            %LET dsid=%SYSFUNC(OPEN(&DATA,i));
            %LET n   =%SYSFUNC(VARNUM(&dsid,&BYVAR));
            %LET fmt =%SYSFUNC(VARFMT(&dsid,&n));
            %PUT &fmt;
            %LET rc=%SYSFUNC(CLOSE(&dsid));
```

```
                    %END;

DATA FINAL&LOG1;
MERGE RESULT&LOG1 CJRR&LOG1 TFR&LOG1;
 BY &LOG3;

  IF TFRSRSSE^=0 THEN DEFT = TFRCSE/(TFRSRSSE);
ELSE DEFT=.;

IF ASFR1SRSSE^=0 THEN ASFR1DEFT = ASFR1CSE/(ASFR1SRSSE);
ELSE ASFR1DEFT=.;
IF ASFR2SRSSE^=0 THEN ASFR2DEFT = ASFR2CSE/(ASFR2SRSSE);
ELSE ASFR2DEFT=.;
IF ASFR3SRSSE^=0 THEN ASFR3DEFT = ASFR3CSE/(ASFR3SRSSE);
ELSE ASFR3DEFT=.;
IF ASFR4SRSSE^=0 THEN ASFR4DEFT = ASFR4CSE/(ASFR4SRSSE);
ELSE ASFR4DEFT=.;
IF ASFR5SRSSE^=0 THEN ASFR5DEFT = ASFR5CSE/(ASFR5SRSSE);
ELSE ASFR5DEFT=.;
IF ASFR6SRSSE^=0 THEN ASFR6DEFT = ASFR6CSE/(ASFR6SRSSE);
ELSE ASFR6DEFT=.;
IF ASFR7SRSSE^=0 THEN ASFR7DEFT = ASFR7CSE/(ASFR7SRSSE);
ELSE ASFR7DEFT=.;

 IF TFR^=0 THEN RELERROR = TFRCSE/TFR;
ELSE RELERROR=.;

 IF ASFR1519^=0 THEN ASFR1RE = ASFR1CSE/ASFR1519;
ELSE ASFR1RE=.;
 IF ASFR2024^=0 THEN ASFR2RE = ASFR2CSE/ASFR2024;
ELSE ASFR2RE=.;
 IF ASFR2529^=0 THEN ASFR3RE = ASFR3CSE/ASFR2529;
ELSE ASFR3RE=.;
 IF ASFR3034^=0 THEN ASFR4RE = ASFR4CSE/ASFR3034;
ELSE ASFR4RE=.;
 IF ASFR3539^=0 THEN ASFR5RE = ASFR5CSE/ASFR3539;
ELSE ASFR5RE=.;
 IF ASFR4044^=0 THEN ASFR6RE = ASFR6CSE/ASFR4044;
ELSE ASFR6RE=.;
 IF ASFR4549^=0 THEN ASFR7RE = ASFR7CSE/ASFR4549;
ELSE ASFR7RE=.;

 LOWER =TFR -2*TFRCSE;  UPPER= TFR +2*TFRCSE;
 IF LOWER<0 THEN LOWER=0;

 ASFR1L =ASFR1519 -2*ASFR1CSE;  ASFR1U= ASFR1519 +2*ASFR1CSE;
 IF ASFRL1<0 THEN ASFRL1=0;
 ASFR2L =ASFR2024 -2*ASFR2CSE;  ASFR2U= ASFR2024 +2*ASFR2CSE;
 IF ASFRL2<0 THEN ASFRL2=0;
```

```
 ASFR3L =ASFR2529 -2*ASFR3CSE;  ASFR3U= ASFR2529 +2*ASFR3CSE;
 IF ASFRL3<0 THEN ASFRL3=0;
 ASFR4L =ASFR3034 -2*ASFR4CSE;  ASFR4U= ASFR3034 +2*ASFR4CSE;
 IF ASFRL4<0 THEN ASFRL4=0;
 ASFR5L =ASFR3539 -2*ASFR5CSE;  ASFR5U= ASFR3539 +2*ASFR5CSE;
 IF ASFRL5<0 THEN ASFRL5=0;
 ASFR6L =ASFR4044 -2*ASFR6CSE;  ASFR6U= ASFR4044 +2*ASFR6CSE;
 IF ASFRL6<0 THEN ASFRL6=0;
 ASFR7L =ASFR4549 -2*ASFR7CSE;  ASFR7U= ASFR4549 +2*ASFR7CSE;
 IF ASFRL7<0 THEN ASFRL7=0;

 LABEL='TOTAL FERTILITY RATE';

  TYPE='Rate';

 %IF &BYVAR^=   %THEN %DO;
FORMAT &LOG3 &fmt.;
%END;

CLUSTER=&NCLUSTER;
STRATA=&NOSTRATA;
WOMEN=&stop;

KEEP LABEL TYPE TFR TFRCSE W_N TFRSRSSE DEFT RELERROR LOWER UPPER
&LOG3
ASFR1519 ASFR2024 ASFR2529 ASFR3034 ASFR3539 ASFR4044 ASFR4549
ASFR1CSE ASFR2CSE ASFR3CSE ASFR4CSE ASFR5CSE ASFR6CSE ASFR7CSE
 ASFR1DEFT ASFR2DEFT ASFR3DEFT ASFR4DEFT ASFR5DEFT ASFR6DEFT
ASFR7DEFT
ASFR1SRSSE ASFR2SRSSE ASFR3SRSSE ASFR4SRSSE ASFR5SRSSE ASFR6SRSSE
ASFR7SRSSE
ASFR1RE ASFR2RE ASFR3RE ASFR4RE ASFR5RE ASFR6RE ASFR7RE
ASFR1L ASFR2L ASFR3L ASFR4L ASFR5L ASFR6L ASFR7L
ASFR1U ASFR2U ASFR3U ASFR4U ASFR5U ASFR6U ASFR7U
CLUSTER STRATA WOMEN;

RUN;

PROC FORMAT;
VALUE NAF 999999999='NA' 0='Entire sample';
RUN;

OPTIONS LS=120 NONUMBER NODATE FORMDLIM=' ';

PROC SORT DATA=FINAL&LOG1 NODUP;
BY &LOG3;
RUN;

DATA FINAL&LOG1;
```

```
SET FINAL&LOG1;
IF &i=1 THEN &LOG3=0;
IF &i=2 AND &LOG3=. THEN DELETE;
RENAME &LOG3=LEVEL;
RUN;


PROC APPEND BASE=TABLE DATA=FINAL&LOG1;  /*The replicate TFRs in TEMPJRRSs
are appended in the DATAJRR dataset*/
RUN;

%END;

 /*ods rtf file="&intdate &months &byvar &survey";*/

PROC PRINT DATA=TABLE WIDTH=MINIMUM HEADING=HORIZONTAL NOOBS;
 VAR LABEL TYPE TFR TFRCSE LOWER UPPER TFRSRSSE DEFT RELERROR
ASFR1519 ASFR2024 ASFR2529 ASFR3034      ASFR3539 ASFR4044 ASFR4549
ASFR1CSE ASFR2CSE ASFR3CSE ASFR4CSE    ASFR5CSE ASFR6CSE ASFR7CSE
ASFR1L ASFR1U ASFR2L ASFR2U     ASFR3L     ASFR3U     ASFR4L     ASFR4U
        ASFR5L ASFR5U ASFR6L ASFR6U ASFR7L ASFR7U
ASFR1DEFT ASFR2DEFT      ASFR3DEFT ASFR4DEFT      ASFR5DEFT
ASFR6DEFT    ASFR7DEFT
ASFR1SRSSE  ASFR2SRSSE  ASFR3SRSSE  ASFR4SRSSE  ASFR5SRSSE  ASFR6SRSSE
ASFR7SRSSE
ASFR1RE ASFR2RE ASFR3RE ASFR4RE ASFR5RE ASFR6RE ASFR7RE
 CLUSTER STRATA WOMEN W_N;
 BY LEVEL;
  FORMAT LEVEL NAF.;
TITLE "&intdate &months &byvar &survey";
RUN;

/*ods rtf close;*/

%mend;
```

APPENDIX C: SAS Macro for Infant Mortality Rate

```
%MACRO MORT(DATA,PSU,BYVAR,WEIGHT,INTDATE,SURVEY,YEARS);

libname tfr 'C:\Users\tugba\Desktop\data preparation\SAS DATA';

proc datasets lib=work
nolist kill;
quit;
run;

data &data;
set tfr.&data;
caseidc=1000000000*TDHS+100000*v001+1000*v002+10*v003+bidx;
select (TDHS);
when (&survey) output;
otherwise;
end;
run;

      /*  Assigns deaths to age groups and 5 year periods */

Data iter1 iter2; Set &data;
 array agegrp {10} (0 1 3 6 12 24 36 48 60 0);            /* Creates age cohorts */
 array limits {10};                                 /* Creates analysis periods */

                  /* Sets length of anlysis periods in months (must be <= 60 months) */
 period     = 60;
 limits {1} = &intdate - period;          /* Establishes boundries for the analysis periods */
 Do i = 2 To 10;
   limits {i} = limits {i-1} - period;
 End;
 upplim     = &intdate - 1;                          /* Sets upper limit for analysis */
 lowlim = limits{i-1};                          /* Sets lower limit for analysis */

                        /* Selects only those children who were born in the period */
                        /* of analysis and who had died prior to the interview      */
 If lowlim <= B3 <= upplim & B5 = 0;
 months = B7;                   /* Imputed age at death*/

            /* Creates variable agedth which will be age group in which the child died */
 agedth = .;

      /* Assigns children who were under 60 months when they died to a death age group */
      Do i = 1 To 8;
        If (agegrp{i} <= months < agegrp{i+1}) Then Do
          agedth = i;
          i = 10;
        End;
```

```
      End;

/* If agedth <> .;                /* Excludes children who died after 5 years of age */

If agedth ^= . ;

 agei   = B3 + agegrp {agedth};                /* lower limit of age group of death */
 nxtage = B3 + agegrp {agedth+1};                /* upper limit of age group of death */


 /* Establishes the period of birth; ...+1 is to be consistent with the analysis periods */
 perborn = Int((&intdate-1 - B3)/period) + 1;

 limlow = limits(perborn);                     /* Lower limit of the birth period */

      /* Upper limit of the birth period */
 If perborn > 1 Then limupp = limits(perborn-1);  Else limupp = upplim + 1;
 n      = 1;
 iter   = 0;

      /* If the child's age group of death spans one period, */
      /* the death is counted once in that period          */
 If limlow <= B3 & nxtage < limupp Then Do iter = 1; n = 1; End;

      /* If the child's age group at death spans two periods,    */
      /* the death is counted laf in each period               */
 If agei < limupp <= nxtage Then Do;
                              iter = 2;
                              n    = 0.5;
                              If perborn = 1 Then Do; iter = 1; n = 1; End;
                           End;

      /* If the child's age group at death spans period i+1, but */
      /* the child was born is period i, the death is counted in period i+1   */
 If B3 < limupp <= agei Then Do;
                          perborn = perborn - 1; iter = 1; n = 1;
                        End;

colper  = perborn;

ADJ_WGT= n * &WEIGHT;    /* Creates the weight that will be used for the data */

IF 1<=COLPER<=5;              /* analysis period between 1 and 5 */
IF ITER^=0 THEN OUTPUT ITER1;

IF ITER=2 THEN OUTPUT ITER2;
RUN;
```

```
        /* The following section crosstabs the first iteration */

PROC SORT DATA=ITER1; BY &psu secu; RUN; /*adding strata here*/
Proc Freq DATA=ITER1 NOPRINT;
 Tables agedth*colper/NoPercent NoRow NoCol Out=deathsH;
 Weight ADJ_WGT;
 BY &psu secu;
Run;

PROC SORT DATA=ITER1; BY CASEIDC; RUN;
Proc Freq DATA=ITER1 NOPRINT;
 Tables agedth*colper/NoPercent NoRow NoCol Out=deathsH2;
 Weight ADJ_WGT;
 BY CASEIDC;
Run;

        /* The following section crosstabs the second iteration */

Data ITER2; Set ITER2; colper = perborn - 1; RUN;

PROC SORT DATA=ITER2; BY &psu secu; RUN;
Proc Freq DATA=ITER2 noprint;
  Tables agedth*colper/NoPercent NoRow NoCol Out=deathsL;
  Weight ADJ_WGT;
  BY &psu secu;
Run;

PROC SORT DATA=ITER2; BY CASEIDC; RUN;
Proc Freq DATA=ITER2 noprint;
  Tables agedth*colper/NoPercent NoRow NoCol Out=deathsL2;
  Weight ADJ_WGT;
  BY CASEIDC;
Run;

PROC SORT DATA=DEATHSH; BY &PSU SECU AGEDTH COLPER; RUN;
PROC SORT DATA=DEATHSL; BY &PSU SECU AGEDTH COLPER; RUN;

/*EKLIYORUM*/
PROC SORT DATA=DEATHSH2; BY CASEIDC AGEDTH COLPER; RUN;
PROC SORT DATA=DEATHSL2; BY CASEIDC AGEDTH COLPER; RUN;


Data Deaths;
 Merge deathsH(Rename=(count=higdth)) deathsL(Rename=(count=lowdth));
 By &psu secu agedth colper;
 nbdths    = Sum(higdth,lowdth);
 IF &YEARS=5  THEN DO; IF COLPER=1; END;             /* RESTRICT CALCULATION */
 IF &YEARS=10 THEN DO; IF COLPER IN (1,2); END;     /* TO 5 OR 10 YEARS BEFORE
SURVEY */
```

```
Run;

Data Deaths2;
 Merge deathsH2(Rename=(count=higdth)) deathsL2(Rename=(count=lowdth));
 By CASEIDC agedth colper;
 nbdths    = Sum(higdth,lowdth);
 IF &YEARS=5  THEN DO; IF COLPER=1; END;              /* RESTRICT CALCULATION */
 IF &YEARS=10 THEN DO; IF COLPER IN (1,2); END;     /* TO 5 OR 10 YEARS BEFORE
SURVEY */
Run;

PROC SORT DATA=DEATHS; BY &psu secu; RUN;

Proc Freq DATA=DEATHS NOPRINT;
 Tables agedth*colper/NoPercent NoRow NoCol Out=deaths;
 Weight nbdths;
 BY &psu secu;
Run;

/*EKLIYORUM*/
PROC SORT DATA=DEATHS2; BY CASEIDC; RUN;

Proc Freq DATA=DEATHS2 NOPRINT;
 Tables agedth*colper/NoPercent NoRow NoCol Out=deaths2;
 Weight nbdths;
 BY CASEIDC;
Run;

%IF &YEARS=10 %THEN %DO;

            PROC SORT DATA=DEATHS; BY &PSU SECU AGEDTH; RUN;
            PROC MEANS DATA=DEATHS SUM NOPRINT;
              VAR COUNT;
              BY &PSU SECU AGEDTH;
              OUTPUT OUT=DEATHS SUM=;
            RUN;

            DATA DEATHS; SET DEATHS(DROP=_TYPE_ _FREQ_); COLPER=10; RUN;

  PROC SORT DATA=DEATHS2; BY CASEIDC AGEDTH; RUN;
            PROC MEANS DATA=DEATHS2 SUM NOPRINT;
              VAR COUNT;
              BY CASEIDC AGEDTH;
              OUTPUT OUT=DEATHS2 SUM=;
            RUN;

            DATA  DEATHS2;  SET  DEATHS2(DROP=_TYPE_  _FREQ_);  COLPER=10;
RUN;
```

```
            %END;

Data births; Set &data;
 array agegrp {10} (0 1 3 6 12 24 36 48 60 0);           /* Creates age groups */
 array limits {10};                                 /* Creates analysis periods */

     /* Sets length of analysis periods in months (must be <= 60 months) */
 period     = 60;
 limits {1} = &intdate - period;            /* Establishes boundries for the analysis periods */
     Do i = 2 To 10;
      limits {i} = limits {i-1} - period;
     End;

 upplim     = &intdate - 1;                          /* Sets upper limit for analysis */

lowlim = limits{i-1};  /* Sets lower limit for analysis */

     /* Selects only those children who were born in the period of analysis */
If lowlim <= B3 <= upplim;

                             /* Current age or age at death of the child. */
If B5 = 0 Then months = B7; Else months = (&intdate - B3);

                           /* Establishes the period of birth; ...+1 is to */
                          /* be consistent with the analysis periods      */
perborn = Int((&intdate-1 - B3)/period) + 1;

TMPWGT=1;   /**** temporary weight used in calculating unweighted numbers of cases ****/

%IF &BYVAR=  %THEN %LET TEMP_BY=99999;
         %ELSE %LET TEMP_BY=&BYVAR;
 W_BYVAR=&TEMP_BY;

/* The following section tabulates exposure in the different age groups */
/*  ageexp is equal to the number of the age group being tabulated   */

%MACRO EXPOSURE(VARBY=, WEIGHT2=, OUTDSN=);

%DO i=1 %TO 8;

Data iter1 iter2; Set Births;
 array agegrp {10} agegrp1-agegrp10;
 array limits {10} limits1-limits10;
 ageexp = &i;     /* Sets the number of the current age group */
 IF PERBORN>0;

     /* Sets the lower bound for the period in which the child was born */

 limlow = limits(perborn);
```

```
      /* Sets the upper bound for the period in which the child was born */

 If perborn > 1 Then limupp = limits(perborn-1); Else limupp = upplim + 1;
                               /* Sets the lower bound for the child's age group */
 agei   = B3 + agegrp {ageexp};
 nxtage = B3 + agegrp {ageexp+1};   /* Sets the upper bound for the child's age group */

      /* Selects children exposed for at least part of the age */
      /* group; i.e. children who enter the age group          */
 If agegrp {ageexp} <= months;
 iter   = 0;

      /* All exposure occurs in period following birth period */
      If limupp <= agei Then Do;
        perborn = perborn - 1; iter = 1; n = 1;
        If perborn > 0 Then limlow = limits(perborn);
        If perborn > 1 Then limupp = limits(perborn-1);
                     Else limupp = upplim + 1;
      End;
      /* All exposure occurs in birth period */
      If nxtage < limupp Then Do iter = 1; n = 1; End;

      /* Exposure occurs in period of birth and following period  */
      /* child is counted as half in both periods                 */
      If agei < limupp <= nxtage Then Do;
        iter = 2;
        n    = 0.5;
        If perborn = 1 Then iter = 1;
      End;
      colper  = perborn;

 ADJ_WGT = n * &WEIGHT;    /* Creates a weight for the data */

IF 1<=COLPER<=5;
IF ITER^=0 THEN OUTPUT ITER1;
IF ITER=2 THEN OUTPUT ITER2;
RUN;
        /* Tabluates the first part of the exposure in the age group by period of analysis   */

PROC SORT DATA=ITER1; BY &VARBY; RUN;
      Proc Freq DATA=ITER1 noprint;
           Tables ageexp*colper/NoPercent NoRow NoCol Out=ExposH;
           Weight &weight2;
           BY &VARBY;
      Run;

Data ITER2; SET ITER2;  colper = perborn - 1; RUN;
```

```
    /* Tabulates the second part of the exposure in the age group by period of analysis    */

PROC SORT DATA=ITER2; BY &VARBY; RUN;

    Proc Freq DATA=ITER2 noprint;
       Tables ageexp*colper/NoPercent NoRow NoCol Out=ExposL;
       Weight &weight2;
       BY &VARBY;
    Run;

PROC SORT DATA=EXPOSH; BY &VARBY AGEEXP COLPER; RUN;
PROC SORT DATA=EXPOSL; BY &VARBY AGEEXP COLPER; RUN;

Data BirthsE;
    Merge ExposH(Rename=(count=higexp)) ExposL(Rename=(count=lowexp));
        By &VARBY ageexp colper;

    nbchild   = Sum(higexp,lowexp);  /* Creates a weight for the data */

 /* Tabulates exposure in the age group by period of analysis  and writes results to a file called
exposi  */

 IF &YEARS=5  THEN DO; IF COLPER=1; END;         /* RESTRICT CALCULATION */
 IF &YEARS=10 THEN DO; IF COLPER IN (1,2); END;   /* TO 5 OR 10 YEARS BEFORE
SURVEY */
RUN;

PROC SORT DATA=BIRTHSE; BY &VARBY; RUN;

    Proc Freq DATA=BIRTHSE noprint;
       Tables ageexp*colper/NoPercent NoRow NoCol Out=Expos&i;
       Weight nbchild;
       BY &VARBY;
    Run;

PROC SORT DATA=EXPOS&i; BY &VARBY AGEEXP COLPER;
RUN;

%END;

Data EXPO_ALL;
 Merge Expos1 Expos2 Expos3 Expos4 Expos5 Expos6 Expos7 Expos8;
 By &VARBY ageexp colper;
RUN;

PROC SORT DATA=EXPO_ALL; BY &VARBY; RUN;

 Proc Freq DATA=EXPO_ALL NOPRINT;
  Tables ageexp*colper/NoPercent NoRow NoCol Out=&OUTDSN;
```

```
  Weight Count;
  BY &VARBY;
 /* Title 'Exposure of children'; */
 Run;

%IF &YEARS=10 %THEN %DO;

                        PROC SORT  DATA=&OUTDSN; BY &VARBY AGEEXP; RUN;
                        PROC MEANS DATA=&OUTDSN SUM NOPRINT;
                           VAR COUNT;
                           BY &VARBY AGEEXP;
                           OUTPUT OUT=&OUTDSN SUM=;
                        RUN;

                        DATA  &OUTDSN;  SET  &OUTDSN(DROP=_TYPE_  _FREQ_);  COLPER=10;
RUN;
                    %END;

%MEND EXPOSURE;

%EXPOSURE(VARBY=&psu secu, WEIGHT2=adj_wgt, OUTDSN=exposure);

PROC SORT DATA=&data OUT=CLUSTER NODUPKEY;
BY &psu secu;
RUN;

DATA CLUSTER;
SET CLUSTER;
 %IF &BYVAR=  %THEN %LET TEMP_BY=99999;
             %ELSE %LET TEMP_BY=&BYVAR;
 W_BYVAR=&TEMP_BY; KEEP &psu secu W_BYVAR;
RUN;

PROC SORT DATA=EXPOSURE; BY &psu secu; RUN;
DATA  EXPOSURE;  MERGE  EXPOSURE(IN=IN1)  CLUSTER(IN=IN2);  BY  &psu  secu;  IF
IN1; RUN;

  /***** Calculating weighted numbers of children exposed to death ****/

Proc Freq DATA=EXPOSURE noprint;
  Tables ageexp*colper/NoPercent NoRow NoCol out=national;
   Weight Count;
Run;

Data national; set national; W_BYVAR=.; run;

PROC SORT DATA=EXPOSURE; BY W_BYVAR; RUN;
Proc Freq DATA=EXPOSURE noprint;
  Tables ageexp*colper/NoPercent NoRow NoCol out=domains;
```

```
  Weight Count;
  by W_BYVAR;
Run;

Data nwgt;
set national domains;
rename count=nwgt;
format count 8.0;
run;

        /***** Calculating unweighted numbers of children exposed to death ****/

%EXPOSURE(VARBY=W_BYVAR, WEIGHT2=TMPWGT, OUTDSN=UNWGT);

%EXPOSURE(VARBY=W_BYVAR CASEIDC, WEIGHT2=ADJ_WGT, OUTDSN=EXPSRS);

Proc Freq DATA=UNWGT noprint;
  Tables ageexp*colper/NoPercent NoRow NoCol out=nationa2;
   Weight Count;
Run;

Data nationa2;
set nationa2;
W_BYVAR=.;
run;

PROC SORT DATA=UNWGT; BY W_BYVAR; RUN;
Proc Freq DATA=UNWGT noprint;
  Tables ageexp*colper/NoPercent NoRow NoCol out=domains2;
  Weight Count;
  by W_BYVAR;
Run;
Data unwgt;
set nationa2 domains2;
rename count=unwgt;
format count 8.0;
run;

PROC SORT DATA=NWGT;  BY W_BYVAR AGEEXP COLPER; RUN;
PROC SORT DATA=UNWGT; BY W_BYVAR AGEEXP COLPER; RUN;

DATA N_ALL;
MERGE NWGT UNWGT;
BY W_BYVAR AGEEXP COLPER;
RUN;

   /* Calculates probabilities of dying and final mortality rates */

PROC SORT DATA=DEATHS;   BY &psu secu AGEDTH COLPER; RUN;
```

```
PROC SORT DATA=EXPOSURE; BY &psu secu AGEEXP COLPER; RUN;

PROC SORT DATA=DEATHS2; BY CASEIDC AGEDTH COLPER; RUN;
PROC SORT DATA=EXPSRS;  BY CASEIDC AGEEXP COLPER; RUN;

Data Probf;
 Merge exposure(Rename=(count=expo  ageexp=ageprb))
       deaths  (Rename=(count=death agedth=ageprb));
 By &psu secu ageprb colper;
 if death=. then death=0;
label AGEPRB='Age in months probabilities'
      COLPER='Five years periods of analysis';
Run;

Data Probf2;
 Merge expSRS(Rename=(count=expo  ageexp=ageprb))
       deaths2  (Rename=(count=death agedth=ageprb));
 By CASEIDC ageprb colper;
 if death=. then death=0;
label AGEPRB='Age in months probabilities'
      COLPER='Five years periods of analysis';
Run;

        /** renumbering clusters if there are incompleted clusters **/

PROC SORT DATA=PROBF;     BY &psu secu; RUN; /*HERE I'M RENUMBERING STRATA*/
 DATA PROBF; SET PROBF; BY &psu secu;
   RETAIN RENUMBERS 0;
   IF FIRST.&psu THEN RENUMBERS=RENUMBERS +1;
 RUN;

PROC MEANS DATA=PROBF MAX NOPRINT;
 VAR RENUMBERS;
 OUTPUT OUT=MAXIMUM MAX=MAX;
RUN;

DATA _NULL_; SET MAXIMUM;
 CALL SYMPUT('NSTRATA',MAX);  /* total number of clusters */ /*I CHANGED NCLUSTER TO NSTRATA*/
RUN;

/*Added for SRS*/
 /** renumbering clusters if there are incompleted clusters **/

PROC SORT DATA=PROBF2;     BY CASEIDC; RUN; /*HERE I'M RENUMBERING STRATA*/
 DATA PROBF2; SET PROBF2; BY CASEIDC;
   RETAIN RENUMBERC 0;
```

```
  IF FIRST.CASEIDC THEN RENUMBERC=RENUMBERC +1; /*I'M CALLING
RENUMBER RENUMBERC FOR CASE*/
 RUN;

PROC MEANS DATA=PROBF2 MAX NOPRINT;
 VAR RENUMBERC;
OUTPUT OUT=MAXIMUM2 MAX=MAX;
RUN;

DATA _NULL_; SET MAXIMUM2;
 CALL SYMPUT('NCASES',MAX);  /* total number of clusters */ /*I CHANGED NCLUSTER
TO NSTRATA*/
RUN;

PROC MEANS DATA=CLUSTER N NOPRINT; /*this creates a strata level data set with a
variable that shows the number of clusters each*/
by &PSU;
VAR secu;
 OUTPUT OUT=NOPSU N=NSECU;
RUN;

DATA PROBF; /*this step adds the number of clusters information to the cluster level data*/
MERGE PROBF NOPSU;
BY &PSU;
RUN;

/*to find the number of PSUs per stratum/
data _null_;
set nopsu;
suffix=put(_n_,5.);
call symput(cats('nosecu',suffix), NSECU);
run;

%put _user_;

PROC DATASETS LIBRARY=WORK NOLIST;
 SAVE &data PROBF PROBF2 cluster nopsu N_ALL BIRTHS/MT=DATA ;  /*changing birth to
births-misspelled data set name*/
RUN;

QUIT;
%put _user_;
%let nosecu0=1;

        /** Computing the mortality rates using JKK1 **/

%DO j=0 %TO &NSTRATA;
%do i=1 %to &&nosecu&j;
```

```
/* when j=0, no deletion occurs */ /*i changed to strata*/
  DATA TEMP; SET PROBF;
  IF (RENUMBERS=&j and secu=&i) THEN do;                /* deletes one cluster at a time
*/
               %if %eval(&j)>0 %then %do;
                             call symput('ran',W_BYVAR);
                            %end;
               %else %let ran=0;
               DELETE;
             end;
          RUN;

 DATA TEMP; SET TEMP;
  IF (RENUMBERS=&j) THEN do;              /* deletes one cluster at a time */
               expo=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*expo;

death=%sysevalf(&&nosecu&i/(&&nosecu&i-1))*death;
             end;
          RUN;

PROC SUMMARY DATA=TEMP NOPRINT;
 VAR EXPO DEATH;
 CLASS W_BYVAR AGEPRB COLPER;
 OUTPUT OUT=TEMP SUM=;
RUN;

DATA TEMP; SET TEMP(DROP=_TYPE_ _FREQ_); IF AGEPRB^=. AND COLPER^=.;
RUN;

PROC SORT DATA=TEMP; BY COLPER; RUN;   /* The sort of Prob is used to have the  */
                       /* different age cohorts listed sequentially */

Data TEMP; SET TEMP;
   probs = death*1000000/expo; /* 1000000 is used to keep the number of decimals */

   /* Mortrate indicates each mortality rate e.g. 1 = neonatal, 3 = infant, 4 = child, and is set to the
last age group  for each rate */
   mortrate = 99;
   If ageprb = 1 Then mortrate = 1;
   If ageprb = 4 Then mortrate = 3;
   If ageprb = 8 Then mortrate = 4;

 /* First step in calculating neonatal mortality; nm is the probability of surviving the first month of
life       */

   If ageprb = 1 Then Do; probdc = (1000000-probs);
                nm     = probdc;
             End;
```

```
    /* First step  in calculating Infant mortality (1q0); infant is the probability of surviving the first
year of life    */
    l1 = Sum(0,Lag1(probs));/* The sum function is used because there */
    l2 = Sum(0,Lag2(probs));/* might be cases where the previous value */
    l3 = Sum(0,Lag3(probs));/* is not defined. */

    /* l1, l2, and l3 are used to calculate the probability of surviving age groups 3, 2, and 1
respectively                            */
    If ageprb = 4 Then Do; probdc = (1000000-probs) *
                                (1000000-   l1) / 1000000 *
                                (1000000-   l2) / 1000000 *
                                (1000000-   l3) / 1000000;
                       infant = probdc;
                      End;

    /* First step  in calculating child mortality (4q1) */
    l1 = Sum(0,Lag1(probs));
    l2 = Sum(0,Lag2(probs));
    l3 = Sum(0,Lag3(probs));

    /* l1, l2, and l3 are used to calculate the probability of surviving age groups 7, 6, and 5
respectively                            */
    If ageprb = 8 Then probdc = (1000000-probs) *
                             (1000000-   l1) / 1000000 *
                             (1000000-   l2) / 1000000 *
                             (1000000-   l3) / 1000000;

    /* Second step in calculating neonatal, infant and child mortality Computes probability of death
(nqx) from probability of  surviving and writes the results to file probs1                 */
    rate = (1000000 - probdc)/1000;
    If (mortrate = 99) Then rate = 0;

LABEL probs='Probability of death-times 1000000';
RUN;

PROC SORT DATA=TEMP; BY W_BYVAR; RUN;
    Proc Freq DATA=TEMP NOPRINT ;
        Tables colper*mortrate/NoPercent NoRow NoCol Out=probs1;
        Weight rate;
        BY W_BYVAR;
      /*   Title 'Mortality rates per 1000 (1)'; */
    Run;

Data TEMP; Set TEMP;
    mortrate = 99;

 /*Calculates postneonatal mortality (inflated by 1000) as infant mortality (1q0) - neonatal
mortality      */
    nm       = Sum(0,Lag3(nm));
```

```
    /* Mortrate = 2 = postneonatal mortality */
    If ageprb = 4 Then Do; mortrate = 2;
                        rate     = (1000000-probdc) - (1000000-nm);
                    End;

    /* Calculates under five mortality (5q0) using the probability of surviving infanthood (0-1) and
the probability of surviving from exact age 1 to exact age 5 (5q0 is inflated by 1000)      */
    infant   = Sum(0,Lag4(infant));
    /* Mortrate = 5 = under five mortality */
    If ageprb = 8 Then Do; mortrate = 5;
                        rate     = 1000000 - (probdc * infant / 1000000);
                    End;
    If mortrate = 99 Then rate = 0;

/* Calculates postneonatal mortality and under five mortality per 1000 and writes the results to file
probs2    */
    rate = rate / 1000;
RUN;

PROC SORT DATA=TEMP; BY W_BYVAR; RUN;
    Proc Freq DATA=TEMP NOPRINT;
        Tables colper*mortrate/NoPercent NoRow NoCol Out=probs2;
        Weight rate;
        BY W_BYVAR;
      /*   Title 'Mortality rates per 1000 (2)';   */
    Run;

Data TEMP; Set probs1 probs2;
label COLPER='Five years periods of analysis'
     MORTRATE= 'Mortality rate';
RUN;

      /* Creates the mortality table  which contains the mortality rates calculated above        */

 PROC SORT DATA=TEMP; BY W_BYVAR; RUN;
 Proc Freq DATA=TEMP NOPRINT;
 %if &j=0 %then %do; Tables colper*mortrate/NoPercent NoRow NoCol OUT=TEMP&j; %end;
 %if  &j>0  %then  %do;  Tables  colper*mortrate/NoPercent  NoRow  NoCol  OUT=TEMP&j&i;
%end;
  Weight Count;
  BY W_BYVAR;
 /* Title 'Mortality rates per 1000'; */
 Run;

%if &j=0 %then %do;
data temp&j;
set temp&j;
ran=&ran+0;
run;
```

```
%end;

%else %do; DATA TEMP&j&i; SET TEMP&j&i; RAN=&RAN+0; replicate=&j&i; RUN;
%end;

%IF &j>0 %THEN %DO;
            PROC APPEND BASE=TOTAL DATA=TEMP&j&i; RUN;

 /* TEMP0 contains the rates for national and for differents levels of by variable */

PROC DATASETS LIBRARY=WORK NOLIST;
 SAVE &data /*CHILD*/ PROBF PROBF2 TEMP0 TOTAL cluster nopsu/*UNWGT NWGT*/
N_ALL/MT=DATA;
RUN; QUIT;

          %END;

%END;
%end;


data total;
set total;
calcstrata=floor(replicate/10);
calcstrata10=calcstrata*10;
calcclu=replicate-calcstrata10;
run;

data _null_;
set nopsu;
RN=int(1+NSECU*ranuni(3674));
suffix=put(_n_,5.);
call symput(cats('RN', suffix), RN);
run;

/*reduce total - delete a random cluster*/

data total;
set total;
%do j=1 %to &nstrata;
if (calcstrata=&j and calcclu=&&RN&j) then delete;
%end;
run;

/*JRR for SRS variance*/
options nonotes;

%do J=1 %to &NCASES;*NUMBER OF LINES HERE*/;
```

```
/* when j=0, no deletion occurs */ /*changed to strata*/
   DATA TEMP2; SET PROBF2;
   IF (RENUMBERC=&J) THEN do;                          /* deletes one case at a time */
                                call symput('ran',W_BYVAR);
                                DELETE;
                      end;
                RUN;

PROC SUMMARY DATA=TEMP2 NOPRINT;
 VAR EXPO DEATH;
 CLASS W_BYVAR AGEPRB COLPER;
 OUTPUT OUT=TEMP2 SUM=;
RUN;

DATA TEMP2; SET TEMP2(DROP=_TYPE_ _FREQ_); IF AGEPRB^=. AND COLPER^=.;
RUN;

PROC SORT DATA=TEMP2; BY COLPER; RUN;   /* The sort of Prob is used to have the
different age cohorts listed sequentially */

Data TEMP2; SET TEMP2;
    probs = death*1000000/expo; /* 1000000 is used to keep the number of decimals */

    /* Mortrate indicates each mortality rate e.g. 1 = neonatal, 3 = infant, 4 = child, and is set to the
last age group  for each rate*/
    mortrate = 99;
    If ageprb = 1 Then mortrate = 1;
    If ageprb = 4 Then mortrate = 3;
    If ageprb = 8 Then mortrate = 4;

    /* First step in calculating neonatal mortality; nm is the probability of surviving the first month
of life        */
    If ageprb = 1 Then Do; probdc = (1000000-probs);
                         nm     = probdc;
                      End;

    /* First step  in calculating Infant mortality (1q0); infant is the probability of surviving the first
year of life    */
    l1 = Sum(0,Lag1(probs));/* The sum function is used because there */
    l2 = Sum(0,Lag2(probs));/* might be cases where the previous value */
    l3 = Sum(0,Lag3(probs));/* is not defined. */

    /* l1, l2, and l3 are used to calculate the probability of surviving age groups 3, 2, and 1
respectively                         */
    If ageprb = 4 Then Do; probdc = (1000000-probs) *
                              (1000000-  l1) / 1000000 *
                              (1000000-  l2) / 1000000 *
                              (1000000-  l3) / 1000000;
                         infant = probdc;
```

```
                End;

   /* First step  in calculating child mortality (4q1) */
   l1 = Sum(0,Lag1(probs));
   l2 = Sum(0,Lag2(probs));
   l3 = Sum(0,Lag3(probs));

   /* l1, l2, and l3 are used to calculate the probability of surviving age groups 7, 6, and 5
respectively                    */
   If ageprb = 8 Then probdc = (1000000-probs) *
                    (1000000-   l1) / 1000000 *
                    (1000000-   l2) / 1000000 *
                    (1000000-   l3) / 1000000;

   /* Second step in calculating neonatal, infant and child mortality */
   /* Computes probability of death (nqx) from probability of surviving and writes the results to
file probs1            */
   rate = (1000000 - probdc)/1000;
   If (mortrate = 99) Then rate = 0;

LABEL probs='Probability of death-times 1000000';
RUN;

PROC SORT DATA=TEMP2; BY W_BYVAR; RUN;
   Proc Freq DATA=TEMP2 NOPRINT ;
      Tables colper*mortrate/NoPercent NoRow NoCol Out=probs1s;
      Weight rate;
      BY W_BYVAR;
    /*   Title 'Mortality rates per 1000 (1)'; */
   Run;

Data TEMP2; Set TEMP2;
   mortrate = 99;

 /* Calculates postneonatal mortality (inflated by 1000) as infant mortality (1q0) - neonatal
mortality */
   nm       = Sum(0,Lag3(nm));
   /* Mortrate = 2 = postneonatal mortality */
   If ageprb = 4 Then Do; mortrate = 2;
                  rate     = (1000000-probdc) - (1000000-nm);
                End;

   /* Calculates under five mortality (5q0) using the probability    */
   /* of surviving infanthood (0-1) and the probability of surviving */
   /* from exact age 1 to exact age 5 (5q0 is inflated by 1000)      */
   infant   = Sum(0,Lag4(infant));
   /* Mortrate = 5 = under five mortality */
   If ageprb = 8 Then Do; mortrate = 5;
                  rate     = 1000000 - (probdc * infant / 1000000);
```

```
                End;
    If mortrate = 99 Then rate = 0;

    /* Calculates postneonatal mortality and under five mortality */
    /* per 1000 and writes the results to file probs2            */
    rate = rate / 1000;
RUN;

PROC SORT DATA=TEMP2; BY W_BYVAR; RUN;
    Proc Freq DATA=TEMP2 NOPRINT;
        Tables colper*mortrate/NoPercent NoRow NoCol Out=probs2S;
        Weight rate;
        BY W_BYVAR;
      /*   Title 'Mortality rates per 1000 (2)';   */
    Run;

Data TEMP2; Set probs1S probs2S;
label COLPER='Five years periods of analysis'
     MORTRATE= 'Mortality rate';
RUN;
      /* Creates the mortality table  which contains the mortality rates calculated above        */

 PROC SORT DATA=TEMP2; BY W_BYVAR; RUN;
 Proc Freq DATA=TEMP2 NOPRINT;
  Tables colper*mortrate/NoPercent NoRow NoCol OUT=TEMP2&j;
  Weight Count;
  BY W_BYVAR;
 /* Title 'Mortality rates per 1000'; */

 Run;

DATA TEMP2&j; SET TEMP2&j; RAN=&RAN+0; replicate=&j; RUN;

PROC APPEND BASE=TOTALS DATA=TEMP2&j; RUN;
PROC DATASETS LIBRARY=WORK NOLIST;
 SAVE &data /*CHILD*/ PROBF PROBF2 TEMP0 TOTAL TOTALS cluster nopsu /*UNWGT
NWGT*/ N_ALL/MT=DATA; /*unwgt, nwgt deleted already??*/
RUN; QUIT;

%end;
                /* calculating sampling errors */


data _null_;
set probf2;
RN=int(renumberc*ranuni(1));
call symput('RN', RN);
run;
```

```
/*reduce total - delete a random case*/

data totals;
set totals;
if (replicate=&RN) then delete;
%end;
run;

DATA TOTAL2; SET TOTAL;
 IF W_BYVAR^=. AND (W_BYVAR-RAN)^=0 THEN DELETE;
 RR2=COUNT*COUNT;                  /*  R(i) square  */
RUN;

DATA TOTALS2; SET TOTALS;
 IF W_BYVAR^=. AND (W_BYVAR-RAN)^=0 THEN DELETE;
 RR2SRS=COUNT*COUNT;                  /*  R(i) square  */
RUN;

PROC SUMMARY DATA=TOTAL2;
 CLASS  COLPER W_BYVAR MORTRATE;
 VAR RR2 COUNT;
 OUTPUT OUT=RESULT1 SUM=SUM_RR2 SUMCOUNT N(COUNT)=NCLUSTER;
RUN;

PROC SUMMARY DATA=TOTALS2;
 CLASS  COLPER W_BYVAR MORTRATE;
 VAR RR2SRS COUNT;
 OUTPUT OUT=RESULT1S SUM=SUM_RR2SRS SUMCOUNTSRS N(COUNT)=NCASES;
RUN;

DATA RESULT1; SET RESULT1(DROP=_TYPE_ _FREQ_);
 IF MORTRATE=. OR COLPER=. OR W_BYVAR=. THEN DELETE; RUN;

DATA RESULT1S; SET RESULT1S(DROP=_TYPE_ _FREQ_);
 IF MORTRATE=. OR COLPER=. OR W_BYVAR=. THEN DELETE; RUN;

PROC SUMMARY DATA=TOTAL2;
 CLASS  COLPER MORTRATE;
 WHERE W_BYVAR=.;              /* W_BYVAR is missing for replicates at the national level */
 VAR RR2 COUNT;
 OUTPUT OUT=RESULT2 SUM=SUM_RR2 SUMCOUNT N(COUNT)=NCLUSTER;
RUN;

DATA RESULT2; SET RESULT2(DROP=_TYPE_ _FREQ_);
 IF MORTRATE=. OR COLPER=. THEN DELETE;
 W_BYVAR=. ;
RUN;

/*EKLIYORUM*/
```

```
PROC SUMMARY DATA=TOTALS2;
 CLASS  COLPER MORTRATE;
 WHERE W_BYVAR=.;               /* W_BYVAR is missing for replicates at the national level */
 VAR RR2SRS COUNT;
 OUTPUT OUT=RESULT2S SUM=SUM_RR2SRS SUMCOUNTSRS N(COUNT)=NCASES;
RUN;

DATA RESULT2S; SET RESULT2S(DROP=_TYPE_ _FREQ_);
 IF MORTRATE=. OR COLPER=. THEN DELETE;
 W_BYVAR=. ;
RUN;

DATA RESULT; SET RESULT1 RESULT2; RUN;
DATA RESULTS; SET RESULT1S RESULT2S; RUN;

DATA TEMP0; SET TEMP0; RENAME COUNT=RATE; RUN;

PROC SORT DATA=N_ALL; BY W_BYVAR; RUN;

  PROC MEANS DATA=N_ALL NOPRINT; VAR NWGT UNWGT; BY W_BYVAR;
  WHERE AGEEXP=1; OUTPUT OUT=MORT1 MIN=N_WGT N_UNWGT; RUN;
  DATA  MORT1;  SET  MORT1(KEEP=W_BYVAR  N_WGT  N_UNWGT);  MORTRATE=1;
RUN;

  PROC MEANS DATA=N_ALL NOPRINT; VAR NWGT UNWGT; BY W_BYVAR;
  WHERE AGEEXP IN (1,2,3,4); OUTPUT OUT=MORT2 MIN=N_WGT N_UNWGT; RUN;
  DATA  MORT2;  SET  MORT2(KEEP=W_BYVAR  N_WGT  N_UNWGT);  MORTRATE=2;
RUN;
  PROC MEANS DATA=N_ALL NOPRINT; VAR NWGT UNWGT; BY W_BYVAR;
  WHERE AGEEXP IN (1,2,3,4); OUTPUT OUT=MORT3 MIN=N_WGT N_UNWGT; RUN;
  DATA  MORT3;  SET  MORT3(KEEP=W_BYVAR  N_WGT  N_UNWGT);  MORTRATE=3;
RUN;

  PROC MEANS DATA=N_ALL NOPRINT; VAR NWGT UNWGT; BY W_BYVAR;
  WHERE AGEEXP IN (5,6,7,8); OUTPUT OUT=MORT4 MIN=N_WGT N_UNWGT; RUN;
  DATA  MORT4;  SET  MORT4(KEEP=W_BYVAR  N_WGT  N_UNWGT);  MORTRATE=4;
RUN;

  PROC MEANS DATA=N_ALL NOPRINT; VAR NWGT UNWGT; BY W_BYVAR;
   OUTPUT OUT=MORT5 MIN=N_WGT N_UNWGT; RUN;
  DATA  MORT5;  SET  MORT5(KEEP=W_BYVAR  N_WGT  N_UNWGT);  MORTRATE=5;
RUN;

DATA N_ALL; SET MORT1 MORT2 MORT3 MORT4 MORT5; RUN;

PROC SORT DATA=N_ALL;  BY W_BYVAR MORTRATE; RUN;
PROC SORT DATA=TEMP0; BY W_BYVAR MORTRATE; RUN;

DATA TEMP0; MERGE TEMP0 N_ALL; BY W_BYVAR MORTRATE; RUN;
```

```
               /*** combining all info ****/

PROC SORT DATA=RESULT; BY COLPER W_BYVAR MORTRATE; RUN;
PROC SORT DATA=TEMP0;  BY COLPER W_BYVAR MORTRATE; RUN;

PROC SORT DATA=RESULTS;  BY COLPER W_BYVAR MORTRATE; RUN;

PROC FORMAT;
 VALUE MORTF 1='NEONATAL'   2='POSTNEONATAL'    3='INFANT'
            4='CHILD'      5='UNDER 5';
RUN;
               %IF &BYVAR^= %THEN %DO;
                                      %LET dsid =%SYSFUNC(OPEN(&data,i));
                                      %LET n    =%SYSFUNC(VARNUM(&dsid,&BYVAR));
                                      %LET fmt  =%SYSFUNC(VARFMT(&dsid,&n));
                                      %LET rc   =%SYSFUNC(CLOSE(&dsid));
                                 %END;

DATA ALL; MERGE TEMP0 RESULT RESULTS;
 LENGTH VARIABLE $ 8;
 BY COLPER W_BYVAR MORTRATE;
VARIANCE=(SUM_RR2                    -                    2*SUMCOUNT*RATE                    +
NCLUSTER*RATE*RATE)*(NCLUSTER-1)/NCLUSTER;
SRSVAR=(SUM_RR2SRS-2*SUMCOUNTSRS*RATE+NCASES*RATE*RATE);

SRSSE=(SRSVAR)**0.5;
NEWDEFT=SQRT(VARIANCE)/SRSSE;
STDERROR=sqrt(VARIANCE);
 IF RATE^=0 THEN RELERROR = STDERROR/RATE; ELSE RELERROR=.;
 LOWER =RATE -2*STDERROR;  UPPER= RATE +2*STDERROR;
 IF LOWER<0 THEN LOWER=0;
 TYPE='Rate';
 VAR_SRS=RATE*(1000-RATE)/(N_UNWGT*1000000);
 IF VAR_SRS^=0 THEN DEFT=SQRT(VARIANCE/VAR_SRS)/1000; ELSE DEFT=.;
 SRS=SQRT(VAR_SRS)*1000;
 IF MORTRATE=1 THEN VARIABLE='NEOMORT';
 IF MORTRATE=2 THEN VARIABLE='PNMORT ';
 IF MORTRATE=3 THEN VARIABLE='INMORT ';
 IF MORTRATE=4 THEN VARIABLE='CMORT  ';
 IF MORTRATE=5 THEN VARIABLE='U5MORT ';
 FORMAT MORTRATE MORTF. N_UNWGT N_WGT 8.0;
 FORMAT RATE STDERROR RELERROR LOWER UPPER SRS DEFT 8.3;

 %IF  &BYVAR^=  %THEN  %DO;  RENAME  W_BYVAR=&BYVAR;  FORMAT  W_BYVAR
&fmt.; %END;
RUN;

OPTIONS LS=120 NODATE NONUMBER FORMDLIM=' ';
```

```
 %IF &BYVAR=  %THEN %DO;
 PROC PRINT DATA=ALL NOOBS HEADING=HORIZONTAL WIDTH=MINIMUM;
 VAR VARIABLE MORTRATE TYPE RATE STDERROR N_UNWGT N_WGT SRS DEFT
RELERROR LOWER UPPER SRSSE NEWDEFT;
 WHERE W_BYVAR=.;
 TITLE1 "&intdate survey=&survey &years year";
 TITLE3 ' ENTIRE SAMPLE ';
RUN;
                    %END;

%IF &BYVAR^= %THEN %DO;
   PROC SORT DATA=ALL; BY &BYVAR; RUN;
   PROC PRINT DATA=ALL NOOBS HEADING=HORIZONTAL WIDTH=MINIMUM;
     VAR VARIABLE MORTRATE TYPE RATE STDERROR N_UNWGT N_WGT SRS DEFT
RELERROR LOWER UPPER SRSSE NEWDEFT;
     BY &BYVAR; TITLE "&intdate survey=&survey &years year";;
   RUN;

                  %END;

%MEND MORT;
```

APPENDIX D: Infant mortality rate computed for twelve regions from TDHS-2003 and TDHS-2008

		IMR	SE(IMR)	DEFT	CV
Istanbul	1998	19.20	6.378	1.078	0.33
	1998 and 2003	14.87	4.440	1.019	0.30
West Marmara	1998	27.73	15.189	1.118	0.55
	1998 and 2003	25.37	10.293	1.093	0.41
Aegean	1998	22.19	16.277	1.499	0.73
	1998 and 2003	25.98	9.818	1.274	0.38
East Marmara	1998	31.44	11.652	1.006	0.37
	1998 and 2003	33.99	8.474	1.026	0.25
West Anatolia	1998	7.67	5.182	0.875	0.68
	1998 and 2003	14.90	4.885	1.026	0.25
Mediterranean	1998	26.74	7.854	1.196	0.29
	1998 and 2003	31.94	6.018	1.147	0.19
Central Anatolia	1998	38.02	14.518	1.161	0.38
	1998 and 2003	31.33	9.487	1.204	0.30
West Black Sea	1998	25.53	9.995	0.986	0.39
	1998 and 2003	27.63	7.750	1.073	0.28
East Black Sea	1998	30.20	17.664	1.528	0.59
	1998 and 2003	30.32	12.686	1.511	0.42
North East Anatolia	1998	53.73	13.738	1.199	0.26
	1998 and 2003	54.00	7.659	1.001	0.14
Central East Anatolia	1998	41.13	12.516	1.330	0.30
	1998 and 2003	55.53	6.988	0.943	0.13
Southeast Anatolia	1998	38.03	5.927	1.043	0.16
	1998 and 2003	40.19	4.745	1.072	0.12