# ESTIMATING TRENDS IN FERTILITY IN KENYA FROM NON BIRTH HISTORY DATA

B Y

NGUGI PAUL WAWERU

Q56/7314/2017

A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF SCIENCE IN POPULATION STUDIES AT THE POPULATION STUDIES AND RESEARCH INSTITUTE OF THE UNIVERSITY OF NAIROBI

**NOVEMBER 2019**

# Declaration

This research project is my original work and to the best of my knowledge has not been submitted either wholly or partially, to this or any other university for the award of a degree.

Signed: ________________________________ Date:______________________

**Ngugi Paul Waweru**

Q56/7314/2017

This research project has been submitted for examination with our approval as university supervisors:

Signed: ________________________________Date: ______________________

**Prof. Alfred O. Agwanda**

Signed: ________________________________Date: ______________________

**Dr. Anne Khasakhala**

# Dedication

This work is dedicated to my mum, Grace and my late dad, Stephen, for teaching me the value of education at an early age.

# Acknowledgments

First is to thank the Almighty God for the gift of life and blessings to undertake my studies. This research project would not have been successful without the support of people around me. I would like to thank my wife Rose for her personal support and patience during my study. Special thanks go to my children Tiphanie, Trevor and Troy for being understanding during the time I was away from home while studying.

This project would not have been possible without the effort, input, insights, support and comments from my principle supervisor, Prof. Alfred Agwanda. My appreciation also goes to Dr. Anne Khasakhala, my second supervisor, for her dedication, comments, believing in me and encouragement throughout my academic program.

Special mention goes to my colleague and friend, Mr Ngugi Mwenda who introduced me to **TeXstudio** and **JabRef,** thus making the editing of my project easy, am extremely grateful.

Last but not least, I pass my gratitude to other lecturers at Population Studies and Research Institute (PSRI), fellow students and colleagues at KNBS for their encouragement and support during my studies.

# Abstract

This study aimed at determining the extent to which methods for estimating trends in fertility without use of birth history could be used on Kenyan surveys data by employing the own-children and reverse survival methods in estimating fertility trend in the country. This study sought to check if the estimates from OCM and reverse survival were consistent.

The study used data from 2015/16 KIHBS and 2014 KDHS. Data evaluation was done in order to obtain optimal fertility estimates. When the data was examined using age ratios, there was no systematic overstating and understating of age for almost all the reproductive-aged women interviewed during the survey. 2015/16 KIHBS data reported a Whipples index of 49.0 and 57.5 for terminal digits 0 and 5 respectively. Myer's blended index was 2.9 and this was an indication that in general the data was accurate and therefore did not require any adjustment to improve its quality before use. The data had some 17 year old women who have given birth to more than 7 children.

Results from 2015/16 KIHBS showed that reverse survival estimated TFR to be 3.5 as compared to OCM that estimated it to be 3.8. To check for the consistency of the results, 2014 KDHS was used. The results from 2014 KDHS dataset were consistent when using both reverse survival and own children method and also consistent with figures obtained from birth history method. The two indirect methods can give consistent fertility estimates when the reference period is closer to the survey period but in the fourth and fifth year reverse survival method tends to systematically overstate fertility as compared to OCM and therefore the latter is preferred since it can give reliable and consistent current estimates and trends.

This study found out that in the absence of full birth history data, reverse survival and OCM methods can reliably estimate consistent fertility estimates and trend. The study recommends that the Government should embrace the use of indirect techniques that utilize non birth history data to estimate fertility.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1
# Introduction

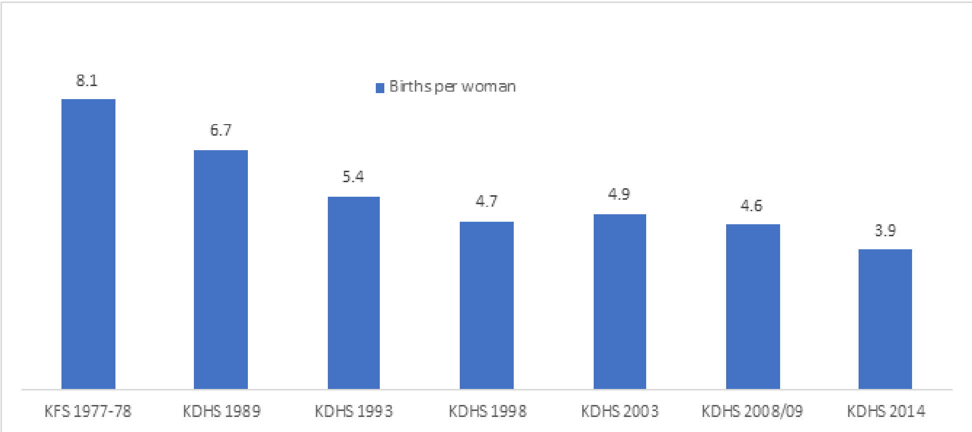
## 1.1 Background to the study

In a country, the composition, size and structure of a population depends on fertility, mortality and migration which are the components of population change. Fertility is known to have effects in the consumption of goods and services which include food, health, education, and housing. Malthus hypothesized that failure to control fertility would lead to high population growth and eventually reduce per capita income which might lead to a reduced consumption below subsistence level. According to Caldwell and Caldwell (1987), population pessimist hold a modern version of this perspective with the argument that if population increases rapidly then it might overwhelm and impact negatively the growth of technological change and capital accumulation. However, according to Simon (1981) and Boserup (1981), population optimists oppose this view by suggesting that on the contrary, population growth leads to societies being creative and innovative in order to keep up with rising consumption demands since economies of scale are supported by larger populations.

According to Kinfu (1993),most sub-Saharan Africa countries view the growth of their population as a problem as well as an endless hindrance to their prospect for future development. The increasing number of countries coming up with population policies that are integrated in their development policy signifies this. The Kenya government has not been left behind in realizing that there is need to manage population as it plays a crucial role in order to actualize economic development that is sustainable as it adopted a national family planning programme in 1967 making it to be among the 1st countries in Sub Saharan Africa to adopt it (NCPD, 2000). The Kenya government has continued to review and come up with better population policies, programmes and strategies aimed at addressing the challenges emanating from population management. For instance, in July 2012, the government adopted a population policy for national development. According to KenyaVision.2030 (2007), the aim of the policy was to ensure that the achievement of the economic, social and political pillars as depicted in Vision 2030's are not hindered by population growth. According to the population policy, during the 1999-2009 inter-censal period, population grew at 2.9 percent and if this growth is not regulated then it would impede the achievement of transforming Kenya in joining the middle-income countries as envisioned by the Vision 2030.

Casterline et al. (2017) observed that about four decades ago, Kenya had one of the highest fertility levels in the world, with Total Fertility Rate (TFR) being reported at 8.1 births per woman . Figure 1.1 shows that over the years there has been notable decline in fertility from 5.4 in 1993 to 4.6 births per woman in 2009. According to Kenya Demographic and Health Survey KDHS (2014), by 2014, TFR for the three years preceding the Kenya Demographic and Health Survey was 3.9 births for each woman, with rural women reporting at least one extra child compared to urban women.

Figure 1.1: Total fertility rate trends, 1978-2014




* Data from 2003 and later are nationally representative while data before 2003 exclude North Eastern region and several northern districts in the Eastern and Rift Valley regions.

Source:KNBS, 2015

Like other developing countries, Kenya does not have a vital registration system that is reliable which leaves surveys and censuses to provide fertility and child mortality estimates. In order to get estimates, various demographic techniques have to be applied. The type of data collected determines if direct or indirect techniques will be used to estimate measures of fertility and mortality. In the 1970, William Brass when interrogating demography data of sub-Saharan Africa countries pioneered indirect techniques of estimating fertility and mortality (United Nations, 1983). In Kenya , the main source of information on fertility rate is data about children ever born (CEB) alive and births that have occurred in a household in the 12 months preceding enumeration obtained from Demographic and Health Surveys (KDHS) and the country's population census conducted every 10 years. The only available option of deducing past fertility trend from the available data is basically through comparison of the rates for the various survey/census years. Nevertheless, when doing such comparison there is a chance of having problems given the possibility of differential coverage and content errors between the different data sources.

Furthermore, it has been argued that the fertility rates reported from such sources may suffer from reference period and memory lapse errors. Based on this argument, the estimates have to be adjusted using indirect techniques many of which often operate under a number of assumptions on past fertility level and pattern. The study of fertility trend based on these indirectly adjusted rates therefore becomes inadequate and suspect. The ideal procedure of reconstructing past fertility trend for countries without dependable vital registration system is by employing birth history data of the type collected in the DHS surveys(Kinfu, 1993). Besides KDHS, other surveys in Kenya do not employ birth history method when collecting data on fertility. Due to unavailability of such data, the Own-Children-Method (OCM) and Reverse Survival provide an alternative indirect procedure of reconstructing the fertility experience of women up to a maximum period of 15 years preceding the enumeration. OCM makes use of information which is invariably obtained in a census or household survey with respect to age and sex of respondents, relationship among members of households and information on children who are enumerated with their mothers. Similarly, reverse survival estimates fertility based on the population data collected from a census or a single-round survey as distributed by sex and age.

### 1.2 Problem Statement

In many developing countries, surveys and censuses suffer from data deficiency and defectiveness (Hlabana, 2010). According to United Nations (2014), some essentials standards have been set for a civil registration system to generate accurate, regular and reliable vital statistics. A key challenge experienced in Kenya according to WHO, (2013) report is that only 57% of all births and 49% of all deaths were being registered and therefore incomplete to provide national or sub national estimates.

Due to the incompleteness of the vital registration systems,information from surveys and censuses are still the main source of fertility data in Kenya. Despite the efforts made to collect complete, accurate and reliable information on fertility, errors due to content and coverage affect the available data. Some of the content errors occur when younger women tend to over-report live births, older women under report live births, age misreporting and wrong dating of births and marital status (Potter, 1977) and (Hlabana, 2010). Coverage errors result from double counting, omission of enumeration area units or population sub-groups .

Coale and Trussell (1996) noted that different demographic techniques have been developed to help in generating fertility estimates depending on the nature of data. It is worthy noting that these indirect methods do not provide foolproof solutions whenever population data is of poor quality (Coale and Trussell, 1996). For results to be reliable, there is need to ensure good accuracy when collecting the data and level of assumptions required by the models should be met by population. So to apply these models there is need to exercise serious prudence.

When vital registration system have incomplete data, fertility rates are usually estimated through the use of data collected from households, either by depending on the age distribution information gathered from the population or by asking women about their childbearing patterns. In the second category, Full Birth History (FBH) method is the dominant method and women are asked to provide the dates of births for all the live births sex of those children as well as their survival status. This method is also known to require detailed data and hence is only applicable in fairly small samples and requires data that is collected periodically (Avery et al., 2013).

Partial birth histories ask for information on total number of children a woman has ever given birth to and are alive together with those children a woman has given birth to but have died. In the absence of information on full birth history including recent births, analysis need to estimate recent fertility through indirect methods. Examples of indirect estimation technique are reverse survival and own children method.

Fertility estimation methods based on full birth history are expensive and can only be used for surveys such as DHS (Avery et al., 2013). Although census data which relies on partial history can be used for lower level estimates, they are collected every 10 years. Experience from 2009 census also indicated that partial birth history tends to underestimate level of fertility due to under-reporting of births(KNBS, 2012). Own children method can be modified for use when birth history is absent and also for sub national level estimates , however, there are limited published results but they have not been applied to data other than census or DHS (Collins and

Levin, 2008) and (Garenne and McCaa, 2017).

Reverse survival model is one of the indirect method used in the estimation of fertility and is preferred since it is the most economical. This method relies on data on age and sex which is collected in single-round survey and census. It ‘reverse survives’ the people who are dead in the population for different ages with the aim of deriving births that occurred some years ago, through survivor-ship probabilities for both children and adults and using Age Specific Fertility Rates (ASFR) (Spoorenberg, 2014). Reverse survival method bears some similarities with own-children method which is also an indirect technique model for fertility estimation (Cho et al., 1986) and (Dubuc, 2009), though the data requirements for own children method are even lower.

However, own-children technique has advantages over other indirect methods which include non requirement of prior assumptions on past vital rates. It also does not require any direct and detailed information on birth statistics. Moreover, the method has the capability producing fertility estimates at various points, hence can give both trends and levels. Such a technique of estimating current and past fertility is very useful for countries like Kenya where specialised fertility surveys are limited and yet reliable information on recent past fertility trends are urgently required. Abbasi-Shavazi (2013) noted that other indirect methods that use partial birth history rely on the assumption that fertility is constant for the period before enumeration. This assumption is likely to bias results for countries still undergoing fertility transition like Kenya.

The Kenya Integrated Household Budget Survey (KIHBS) conducted in 2015 and 2016 collected partial information (partial birth history) that could allow for estimation of recent fertility estimation (TFR) but not trends. The major question is that can KIHBS data be used to produce trends in fertility levels? If so, how do these results compare with those obtained from the birth history? The answer to these questions offered opportunity to extend the methodology beyond survey for two reasons.

(i) Can be used to measures fertility at sub national levels without too much demand on information on birth history which is too expensive.

(ii) Can be modified and used with vital registration data.

Thus in absence of information on complete birth history including recent births, analysis needed to estimate recent fertility through indirect methods. It was within this context that this study made use of the own-children and reverse survival methods in estimating fertility trend in the country. The results from the two methods were used to check for consistency of fertility estimates from previous surveys such as demographic and health surveys (DHS). This study sought to find if the estimates from OCM and reverse survival were consistent.

## 1.3 Research questions

(i) Can data from KIHBS be reliably used to estimate trends in fertility levels in Kenya?

(ii) How do estimates based on OCM and reverse survival compare with those from the Birth history?

## 1.4 Objectives of the study

The main objective was to establish the suitability of OCM and reverse survival to estimate trends in fertility levels from surveys data.

The specific objectives were:

(i) To determine quality of KIHBS data.

(ii) To establish the extent to which data from KIHBS is suitable for estimation of trends in fertility levels.

(iii) To reconstruct trends in fertility levels based on OCM and reverse survival methods.

(iv) To compare trends from non birth history methods with those from DHS.

## 1.5 Justification of the study

The study of fertility is critical in the analysis of demographics since births and deaths are the main factors of population growth. Fertility studies aim in checking and keeping track the advancement made by the available population interventions and programs. Total Fertility Rates (TFR) is ranked among the crucial indices in fertility which help in assessing and evaluating the effects of population programs hence making it's measurement crucial.

Fertility estimations can be done through vital registration system, direct or indirect estimation techniques. In the absence of vital registration system, direct estimation can be done through birth history methods which is common with DHS data sets. Data from FBH are expensive to obtain because they require large samples and longer time to obtain from individual interviews leading to fatigues which sometimes lead to low response rates (Garenne and McCaa, 2017).

Though partial birth history is widely used in countries with incomplete birth registrations it is also known to under-estimate fertility (Avery et al., 2013). It cannot be reliably be used to estimate trends based on the fact that information is at times provided by the head of the household or any other responsible member of the household and may not be privy to all the information about the woman.

The advantage of OCM and reverse survival is that they require fewer data points for estimation and can be used to generate data even in small areas (Avery et al., 2013)and (Childs, 2019). They are also low cost because of data requirements (Avery et al., 2013) and (Collins and Levin, 2008). Reverse survival model relies on data on age and sex which is collected in single-round survey and census. It antecedently back-projects the people who are dead in the population for different ages with the aim of deriving births that occurred some years ago, through survivor-ship probabilities for both children and adults and using Age Specific Fertility Rates (ASFR) (Spoorenberg, 2014). On the other hand, OCM has the capability producing fertility estimates at various points, hence can give both trends and levels. Such a technique of estimating current and past fertility is very useful for countries like Kenya where specialised fertility surveys are limited and yet reliable information on recent past fertility trends are urgently required.

This study argued that it is essential for demographic data to be analysed and interpreted only after evaluating the quality of data first. Also, there is need to determine if the use of chosen demographic technique will provide plausible results. The study endeavored to help in enriching the already available information needed in documentation of fertility in Kenya. The reliability and usefulness of demographic estimates as well as parameters that are derived from surveys and censuses relies on the quality of data, and the method used in the estimation.

### 1.6 Scope and limitations of the Study

This study was national and used data from 2015/16 KIHBS and 2014 KDHS. The former was the first integrated household budget survey to be conducted in Kenya after the introduction of County Governments. The samples for both 2015/16 KIHBS and 2014 KDHS were from the NASSEP V household sampling frame. National representative samples were used with each county getting an independent sample. The samples were distributed in both urban and rural. The 2015/16 KIHBS sample took care of seasonality dividing the households into four quarters. Both surveys were aimed at providing data for various indicators that would help assess advancement made in improving the quality of lives of the population.

2015/16 KIHBS data just like most surveys in Kenya faced a number of challenges that impact on the quality of lifetime fertility data as a result of omission of births as well as problem of recall lapse which is more pronounced in older women. This problem mostly affects dead children, children who are living outside the household and more so if they left the household immediately after birth. Also, if still births, foster children and late foetal deaths are included in the children ever born then this might affect the quality of data.

For 2015/16 KIHBS, the study used the age ratio method, Whipples and the Myer's blended indices and distribution of Children Ever Born (CEB) in the process of evaluating age data. Details on how data was evaluated and assessed are discussed in chapter four under data quality and analysis of research project.

2014 KDHS data uses full birth history to estimate fertility rates but in addition it collects summary birth history data which was used to compare fertility trends generated by OCM and reverse survival methods. This study utilized data from the women individual data file.

# Chapter 2

# Literature Review and Analytical Framework

## 2.1 Introduction

This chapter examines literature on the use of indirect demographic estimates using own children and reverse survival methods with the aim of providing a basis for the study. Section 2.2 begins by presenting existing literature on the need for indirect demographic estimation, section 2.3 presents literature on own children method while section 2.4 presents literature on reverse survival method, section 2.5 presents a summary of the literature review of both own children and reverse survival methods . Section 2.6 presents details on the analytical framework.

## 2.2 Indirect Demographic Estimation

According to Brass and Coale (1969) and United Nations (1983) indirect demographic estimation utilizes a variety of procedures, and entirely utilises information that has been obtained from censuses and single round surveys. It has been clear that data obtained from vital registration systems is found to be deficient and defective and thus even with adjustments, demographic estimates cannot be used, however, results emanating from these direct adjustments can be allowed as long as the estimates are comparable to data from censuses.

Brass and Coale (1969) posits that for demographic estimates from censuses and surveys to be reliable and accurate then coverage and content errors should be small and at minimal level. Therefore, the rationale behind the use of indirect estimation assumes that in case population events for a reference period are not well collected thus making data unreliable, there can be other ways of making the data reliable and accurate by transforming it into expected vital estimates. This idea considers what could probably produce the errors by minimizing and lowering any consequences they could have on the estimated demographic parameters. According to United Nations (1983), indirect demographic methods helps in reducing any inconsistencies and errors found in the data.

Indirect techniques use various procedures to estimate fertility. Some of the procedures use observed variables as long as they are related to a parameter and transforms them into desirable parameter. According to Zlotnik and Hill (1981), with these two procedures, the onus lies in simplifying them with a view to reduce potential variability as compared to the real world. Simplifying age and time period variables is expressed as the cost of using related but observable variables in the estimation procedure.

According to Muhwava and Timæus (1996), assumptions and methodology used in indirect techniques differ based on the demographic patterns of a population. The type of data required determines the choice of the technique adopted as well as the type of assumptions necessary for the model and how robust the model is when treated to conditions which are different from those assumptions.

It is worthy pointing out that when fundamental assumptions are violated, then estimates from indirect techniques will not provide true estimates and hence can lead to errors in estimation(Muhwava and Timæus, 1996). For good understanding of the techniques themselves, it calls for good apprehension of the reasons why systems used to collect data in developing countries are weak and insufficient as well as errors emanating from the demographic data collected from these developing countries.

Previous studies have documented that indirect demographic techniques are necessary when vital events produce incomplete and inaccurate data (Muhwava and Timæus, 1996). That being the case, a questions begs as to why the quality of data collected by developing countries is poor. In the presence of a registration system that is efficient and that can provide size, trends, patterns as well as the outcomes of these vital events, there would be no need to apply indirect techniques. Insufficiency and deficiency of systems of vital registration in developing countries is due to several reasons which include political instability, scarce resources as well as poor and weak institutional management (Cleland, 1996); (United Nations, 1983).

Data from surveys and censuses suffer similar defectiveness and therefore do not provide a full-proof solution to the problems found with the registration of vital events. Shryock and Siegel (1976) and Hlabana (2010) have noted that there are four essentials features of a population census which include simultaneity, individual enumeration, defined periodicity and universality. This essential features make data from censuses to be the next available alternative to vital registration, this is because censuses takes into consideration demographic characteristics for all the people in a given country.

According to United Nations (1983), the quality of data collected from censuses for most developing countries in Africa is poor since most of the countries do not have enough finances and manpower to undertake a credible census. It has been noted that data from these censuses suffer from content and coverage errors since some of the individuals resident in a particular country might not be enumerated during the exercise and also some of the information collected during the enumeration could be erroneous.

Majority of developing countries lack accurate and reliable civil registration system that is required in monitoring fertility rates, levels and trends for a period of time (Hlabana, 2010).

That being the case, fertility rates are usually estimated through the use of data collected from households, either by depending on the age distribution information gathered from the population or by asking women about their childbearing patterns. In the second category, Full Birth History (FBH) method is the dominant method and was majorly used in the 1970s during the implementation of World Fertility Survey, and currently used in the carrying out of the DHS programme (Childs, 2019) and (Avery et al., 2013) and (Collins and Levin, 2008).

In FBH, interviewed women are asked to provide the dates of births for all the live births, sex of those children as well as their survival status. This method is also known to require detailed data and hence is only applicable in fairly small samples and requires data that is collected periodically (Childs, 2019); (Avery et al., 2013) and (Collins and Levin, 2008).

In their paper, ”The Own Children Fertility Estimation Procedure: A Reappraisal”, Avery et al. (2013) examined at the logical basis of both FBH and OCM method with the aim of identifying the source of systematic discrepancies as they estimate fertility. The authors applied FBH and OCM techniques to paired DHS data sets collected in 56 countries. A comparison of TFR for a three year sample period of births from the two methods by simple visualization of the graphs showed that TFR from FBH lie above that generated by OCM in majority of the instances. The authors noted that FBH produces significantly larger TFR as compared to OCM technique, an average difference of 4% in estimates. They also noted that FBH tends to overestimate fertility due to biased sampling.

### 2.3 Own Children Method Application

When making inference for a population using information from the age distribution, own children method (OCM) is the preferred technique since children are linked to their mothers using the mothers line number variable found in the household schedule (Cho et al., 1986). Through the use of this method, a mother's and child's age are used to determine the age of the woman during the time of giving birth to that child and also the time period probably when the same birth took place, hence providing a foundation for calculating age pattern fertility estimates; for children who may not be living with their biological mothers and in some cases where children could have died, then own children method advocates for making fertility level adjustments to the estimates.

Cho et al. (1986) noted that OCM was introduced to use population census data with the aim of generating differential fertility estimates. The method is applicable in any household survey where data on age and sex for everyone in the household has been collected. In additional OCM has the advantage of producing fertility estimates at single years and also for smaller population. OCM has been applied in past studies by Abbasi-Shavazi, (1997) ;Abbasi-Shavazi and McDonald (2002); Dubuc (2009); Coleman and Dubuc, (2010 ).

When studying differential fertility, Cho and Bogue (1970) marked that OCM has three advantages over other conventional methods. Firstly, OCM involves an additional step of coding or matching of biological mothers to their children. Secondly, the method allows the study of differential fertility for a period of 15 years preceding the census or survey period. Lastly, the method allows surveys to be modified hence making the data be available for studying fertility differentials. OCM was noted to be superior to data gotten from CEB as it can provide time associated fertility estimates. The technique can give current fertility rates as compared to CEB data that gives cumulative fertility. Cho and Bogue (1970) argued that over time cumulative fertility suffers from a lag in fertility change as compared to current fertility that gives results of actual changes in recent fertility.

Garenne and McCaa(2017) innovated an excel toolkit called 4 -parameter OCM which was an adaption of the original own children method and does not require use of external data sources when estimating adult mortality. Such appropriate innovations help in derivations of reliable fertility estimates that provide a basis of understanding of spatial variability of fertility regimes in an area and therefore is able to account for spatial disparities in adult mortality.

In the study of "Small Area Estimation of Fertility: Comparing the 4-Parameters Own-Children Method and the Poisson Regression-Based Person-Period Approach", Ndagurwa and Odimegwu (2019) posited that 4-parameters OCM can obtain estimates of female adult mortality based on the survival status of children's mothers which can be derived from a logit transformation of proportions who are maternal orphans,and by indirectly using the logits to estimate adult mortality. They noted that the viability in adult mortality for different geographic areas is reflected by the differences in the derived logits, slopes and intercepts.

OCM has the ability to estimate trends in fertility, Collins and Levin (2008) estimated TFR over time on four Kenyan censuses (1969, 1979, 1989 and 1999). Their results showed that TFR in Kenya increased from about 6 births per woman in 1950s and 1960s to about 8 births during the 1970s before declining to 6.5 births in the late 1980s. TFR was found to have reduced to about 5 births per woman in 1999. They noted that OCM had the strength of providing long term trends that act as external validity checks if more than one accurate data set is available. OCM is more reliable whenever fertility levels, trends and differentials are studied since it provides more revealing results as compared to conventional methods (Collins and Levin, 2008).

While carrying out an assessment of OCM of estimating fertility by birthplace in Australia, Abbasi-Shavazi (1997) pointed out that the technique had been applied to the 1986 data (Jain, 1989) and 1991 census (Dugbaza, 1994) with the aim of studying Aboriginal fertility. Abbasi-Shavazi made a conclusion that Jain and Dugbaza found out the OCM was able to produce reasonable fertility estimates of the Aboriginals.

Also, while examining data from the 2011 Iran census together with 2010 Iran DHS, Abbasi-Shavazi (2013), estimated fertility using OCM technique and compared the results with other direct and indirect methods of estimating fertility. The results showed that OCM was useful in estimating current differential fertility. He noted that other indirect methods rely on the assumption that fertility is constant for the period before enumeration. The assumption would mislead for countries still undergoing fertility transition.Their results showed that yearly fertility fluctuations could only be examined through the use of OCM.

### 2.4 Reverse Survival Application

When estimating for total fertility, reverse survival model is preferred since its estimates are consistent and rarely suffer from any erroneous assumptions emanating from age distribution or if incorrect age patterns, mortality levels, trends are used during the estimation. When using reverse survival method, consistency of the estimates depends on the quality of data for both age and sex variables. When Spoorenberg (2014) was evaluating reverse survival method and how it estimates fertility, he noted that despite the technique being termed as simple and parsimonious to apply it is rarely used.

To show the contribution made by the method in estimating total fertility for past and present trends, Spoorenberg (2014) used data from five countries that included Kenya,Algeria,Ghana, Japan and Mongolia since these countries have distinct data quality issues and fertility levels. An excel template FE_reverse_4.xlsx, that is provided with Timæus and Moultrie (2012) was used to reverse survive a projected population that had been simulated by first projecting for more than 15 years through the use of sets mortality and fertility and also sex and age patterns.

Spoorenberg (2014) noted that reverse survival gets estimates that are closer to original values gotten from the population. He concluded that sensitivity analysis done on the data showed that both incorrect age specific fertility patterns and mortality rates, trends and patterns do not affect reverse survival estimates. Reverse survival suffers from effect of migration and therefore can only be applied at national levels unless there is detailed information about migration.

According to Timæus and Moultrie (2012), reverse survival method can derive fertility estimates just by the shape of the ASFRs without data on birth levels. The technique has the capability of stringing together reverse-survived estimates from successful surveys and census (Spoorenberg, 2014). Some of the limitations with reverse survival method are fertility estimates being biased by child under-reporting in general and mostly for children aged 0 and 1 years. To overcome the bias, fertility estimates can be aggregated by 5 year ages (United Nations, 1983). The other limitation is that the technique uses a constant age specific fertility rate though the shape of the ASFR curve is stable within a 10 year period. Lastly, reverse survival technique does not estimate TFR by social characteristics, for example, education of mother.

### 2.5 Summary of the literature review

When vital registration systems is unreliable,surveys and censuses become the main source of fertility estimates. Majority of the countries have made efforts to collect complete, accurate and reliable information on fertility but errors due to content and coverage still affect these data. Some of the challenges data from surveys and census suffer from include content errors that occur when younger women tend to over-report live births, older women under report live births, age misreporting and wrong dating of births and marital status. Also the data usually face the challenge of coverage errors that arise from double counting, omission of enumeration area units or population sub-groups. Census and survey data usually suffer from age and parity misreporting which can be as a result of age heaping or age exaggeration.(Moultrie and Zaba, 2013). FBH which collects detailed data from women of reproductive age about each live birth they have given birth to and its survival status is preferred by DHS surveys despite the fact they are deemed to be expensive to undertake since they require extensive training and stringent field supervision. FBH does not allow proxy reports and should only be used when the woman to be interviewed is present. This makes it not appropriate for large-scale surveys or censuses.

Spoorenberg (2014) posits that reverse survival gets estimates that are robust and closer to what other methods get. Despite reverse survival technique being economical in terms of required data, own-children method has advantages over other indirect methods due to its non requirement of prior assumptions on past vital rates and also does not require any direct and detailed information on birth statistics.

## 2.6 Analytical framework

In order to estimate fertility, data from 2015/16 KIHBS was used with both own children and reverse survival method. In order to compare trends from non birth history methods with those from birth history, 2014 KDHS was employed. This section discussed assumptions of each method and the required data. Much of the details on the methodology were covered in chapter three of this project.

### 2.6.1 Own Children Method

When data from vital registration system is incomplete, or when key fertility questions have been omitted in surveys and census then own-children method can be used to estimate fertility. This method is used in the estimation of age specific fertility rates for the period before a census or survey through reverse-survival technique. The OCM was originally developed in the 1950s for application with census data to generate estimates of fertility differentials (Avery et al., 2013).

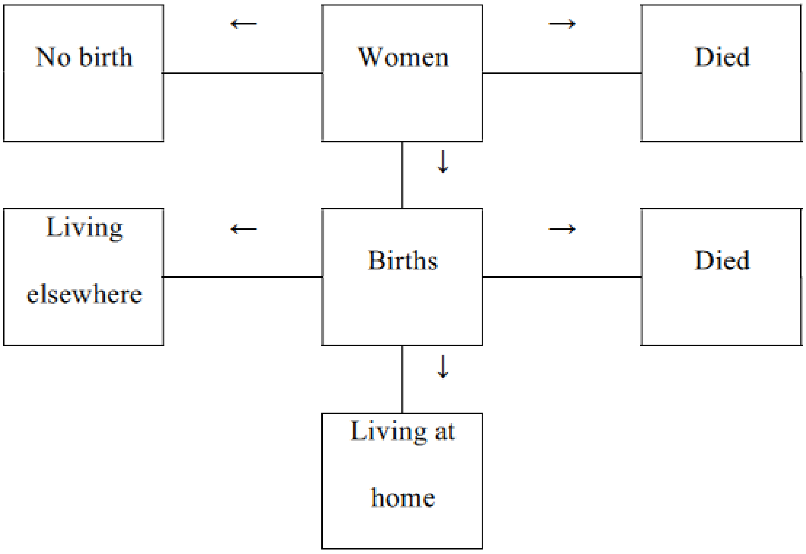
Cho et al. (1986) are credited with developing the first manual for implementing the OCM. The appeal of the OCM technique is that it can be applied by different scholars including those in the field of anthropology who are mostly interested in understanding small communities (Childs, 2019). According to Avery et al. (2013), the OCM can be applied with small samples and any survey data that has age distribution by sex, household indicators namely line number, relationship statuses, and survivorship status of each child's mother.

Moreover, the own-children method provides measures of fertility that are time related and thus are superior to data collected from children-ever born questions. Children ever born data provides cumulative fertility as compared to own children method that provides current fertility estimates. Cho et al. 1970 and Cho et al. (1986) argues that current fertility shows the true change in current fertility rates while there is a fertility lag as time goes by for cumulative fertility.

The 4 parameters OCM technique is direct, and can be represented by Figure 2.1, describing the situation of mothers and children at time of a census/ survey:

In a population, women pass through reproductive ages (from age 12 to 49 years), during

Figure 2.1: Own Children Method Framework




Source:(Garenne and McCaa, 2017)

which they may give birth or not. The women may survive and therefore be present at census, or die, and therefore not counted in the census. If they had births, children may be living with their mother, and therefore appear on the census roster, they may live elsewhere, or may have died. What is enumerated in a census are the surviving women (by age), their biological children living with them, identified by a relationship within the household linking children to their mother and describing mother's line number in the household roster. Furthermore, many surveys and census also collect data on total number of children a woman has ever given birth to in their reproductive ages . In addition, most censuses often provide information on mother's survival, through the question of maternal orphanhood.

The 4-parameters own-children method is simply a way of reconstructing past person-years lived by women as distributed by age in single years and the time period, and past births as cross-tabulated in mothers age in single years and time period from available data, with the aim of computing age-specific fertility rates by time period in the years preceding the census.

Note that for single year of age and period, person- years are equivalent to average population in each cell. Counting children ever-born is especially important in Africa, because a large proportion of children are not living with their mother for example if they are fostered children. Children may be sent elsewhere right after weaning, at time of entry into school, or when they are old enough to do odd jobs. In many African countries the proportion of children living elsewhere is large, and of the same order of magnitude as child mortality. On average, in DHS surveys conducted in sub-Saharan Africa, among women age 15-49, the proportion of children living elsewhere was 19% and the proportion of children who had died was 17%. These figures are obviously very different from the corresponding low values found in North-America or in the Republic of Korea, where the original own children method was tested (Hlabana, 2010).

#### 2.6.1.1 Data needed

To apply the 4-parameters own-children method, the following data are needed, and were readily available in the 2015/16 KIHBS and 2014 KDHS datasets, where children could be linked to their mother

(i) Age structure: number of resident women, by single year of age (12-64). This age group is necessary to calculate the person-years lived in the age group 12-49 in the past 15 years (15 years before the census/survey, women age 64 at census/survey were age 49).

(ii) Fertility: mean children ever born as distributed by single year of age of woman (12-64). In case children ever-born are not available above age 50, one could take the average of children ever-born at age 45-49 for all ages 50-64 years.

(iii) Age of children: number of children living at home with their mother as distributed by age of children in single years (0-49) as well as the distribution of age of mother in single years(12-64). This data can be obtained by matching mother's line number and her children. In addition to distributing births in the past, this table also allows calculating the mean age of children by age of mother, needed to estimate the time spent at risk of mortality after the birth of the child until the census date, and therefore the survival of mothers.

(iv) Female survival: proportion of persons with surviving mother, by single year of age (0-49). This ratio is obtained directly from the orphanhood question (is name's biological mother alive?). If this variable is not available, one could use another source, such as another census, or a DHS survey conducted in the same country at about the same time, or data from another country with similar levels of mortality. By combining the mean age of children tabulated by the age of mother, obtained directly from the preceding table, one can compute the proportion of women who have died in the same cohort since the birth of their children.

The rationale of this calculation is that the survival probability between a birth and the time of the census a cohort of women age 'a' at the time of census, whose children were born on average 'b' years ago, is equal to 1 – the proportion of orphans among children age 'b' at census. This procedure uses data readily available in the census and avoids using model life tables or other approximations.

#### 2.6.1.2 Assumptions

The method is based on a number of assumptions, and in particular on independence between fertility, mortality and migration of children as well as proper age reporting. The assumptions are;

(i) Accurate age of mother;

(ii) Accurate age of children living with mother;

(iii) Accurate matching of children with their biological mother;

(iv) The distribution of dates of birth of children is the same for the three categories of children considered: children living in the household with their mother, children living outside the household or elsewhere, and children who are found dead before census.

As expected, all these assumptions are more or less seriously violated in practice, so that biases in the estimates of fertility rates are always expected. Some of these biases are systematic, in particular violations to 4th assumption.

### 2.6.2 Reverse Survival Method

Reverse survival method is applied during the estimation of fertility rates if data from surveys or census did not have questions that collected fertility directly. If population is assumed to be closed to migration, then individuals aged y years are held as the survivors of the births that happened y completed years antecedently. This means that the number of births that occurred y years ago can be calculated, as long as life table survival probabilities from birth to age y can be estimated, this is denoted by ,Ly/l0. If the population is reverse-survived to its birth year and divided by an estimated total population for that year then crude birth rate is computed. General Fertility Ratio, is computed by dividing the reverse-survived population by an estimate of women of childbearing age. To estimate TFR, we combine estimates of past births and women based on the age as computed from reverse survival models with estimates gotten from age pattern of fertility.

Information regarding births in the last 12 months or last live birth dates only gives an estimate of current fertility while on the other hand, application of reverse survival model gives fertility estimates for the last 15 years. In addition, birth histories provide fertility estimates but only for women aged 15 to 49 years, however, reverse survival provides fertility estimates that are not increasingly truncated at older ages since calculations are done for more distant periods. The only data requirement for the approach to give an annual series of fertility estimates is for single year age distribution for children to be available. Practically, Census and survey age data from developing countries are rarely adequately accurate to yield a time series that is not distorted.

#### 2.6.2.1 Data required

The following data was required to help in the computation of individual year General Fertility Ratios :

(i) Distribution of total individuals aged 0 to 14, by single years of age and sex.

(ii) Distribution of total females aged 15 to 64 years by five year age cohorts.

(iii) Five year age cohorts survival probabilities, Lx, for individuals aged 0 to 14 both males and females.

(iv) Survivorship ratios, $\frac{5Lx-5}{5Lx}$ for individual women for all the three periods of five year preceding the survey/census.

For each of the mortality estimates, the execution of the method in the connected Excel workbook has specification for both reference to period specific parameters $\alpha$ and ß of appropriate relational model life tables, and associated values of children $_5q_o$ and for adult women $_{45}q_{15}$ for the three periods of five years preceding the survey or census.

To compute Total Fertility estimates, the following data was required

(i) One age specific fertility distribution which assumes that the estimates cover the entire period, or

(ii) Two age specific fertility distributions, with the first one covering a date fairly closer to the index under enquiry while the other applying to a date which is roughly 15 years preceding the index.

General and Total Fertility Ratios estimates are almost determined by the estimated number of births in relation to the total women aged 15-49.

#### 2.6.2.2 Assumptions

Reverse survival estimates relies on the premise that population is closed to migration during the reference period. However, since children usually accompany their mothers during migration, then any errors that would arise in both numerator and denominator used in estimating fertility rates are assumed to largely cancel out. Whenever migration flows are deemed significantly large and the population migrating have different fertility rates than the rest of the population then the estimates can suffer from significant bias.

# Chapter 3

# Methodology

## 3.1 Introduction

This chapter gives a description of the data source and methodology to be used during data analysis.

## 3.2 Sources of Data

This study used data from the 2015/16 KIHBS and 2014 KDHS. KIHBS was an integrated household budget survey that was carried out in Kenya but was the first after the promulgation of the Constitution of Kenya which created devolved system of government. Both data sets were nationally representative surveys and were stratified based on the counties.

For KIHBS survey, women aged 12 years and above provided data on fertility and some of the eligible women were interviewed through proxies in case they were not present for interviews. In such cases information about the eligible women was provided by respondents who claim to have the best reproductive knowledge about them. According to Mislevy et al. (1991), accepting proxy answers is known to increase response errors in surveys.

On the other hand, KDHS collects both full birth history and summary birth history data. Therefore despite DHS data being expensive and time consuming to collect, it has the advantage of estimating fertility rates through direct and indirect techniques.

In the 2015/16 KIHBS survey, 25,423 women aged 15-49 years gave information on their lifetime and recent number of births they have had. 2014 KDHS had 31,079 eligible women aged 15-49. Both surveys sought information on fertility by use of the following questions:

(i) How old is (name)?

(ii) Have you ever had a live birth?

(iii) How many children have you borne alive (Male)?

(iv) How many children have you borne alive (Female)?

(v) How many children have you borne alive who usually live in the household (Male)?

(vi) How many children have you borne alive who usually live in the household (Female)?

(vii) How many children have you borne alive who usually live elsewhere (Male)?

(viii) How many children have you borne alive who usually live elsewhere (Female)?

(ix) How many children have you borne alive who have died (Male)?

(x) How many children have you borne alive who have died (Female)?

## 3.3 Methods of data analysis

The data analysis techniques used were, own-children and reverse survival methods. This study used Excel toolkit designed by Garenne and McCaa (2017) to implement a 4-parameters own children method and excel template provided by Timæus and Moultrie (2012) to estimate fertility using reverse survival . The description of the methods were adopted from Garenne and McCaa (2017) as well as Timæus and Moultrie (2012).

### 3.3.1 Own Children Method

When data from vital registration system is incomplete, or when key fertility questions have been omitted in surveys and census then own-children method can be used to estimate fertility.

This method is used in the estimation of age specific fertility rates for the period before a census or survey through reverse-survival technique. The OCM was originally developed in the 1950s for application with census data to generate estimates of fertility differentials (Avery et al., 2013).

Cho et al. (1986) are credited with developing the first manual for implementing the OCM. The appeal of the OCM technique is that it can be applied by different scholars including those in the field of anthropology who are mostly interested in understanding small communities (Childs, 2019). According to Avery et al. (2013), the OCM can be applied with small samples and any survey data that has age distribution by sex, household indicators namely line number, relationship statuses, and survivorship status of each child's mother.

Moreover, the own-children method provides measures of fertility that are time related and thus are superior to data collected from children-ever born questions. Children ever born data provides cumulative fertility as compared to own children method that provides current fertility estimates. (Cho et al. 1970; Cho et al. (1986)) argues that current fertility shows the true change in current fertility rates while there is a fertility lag as time goes by for cumulative fertility.

The study was designed to appraise the Excel toolkit designed by Garenne and McCaa (2017) to implement a 4-parameter own children method. To implement a 4-parameter own children technique, the first step starts by considering a cohort of surviving women and taking into account their mortality since delivery. We then calculate the mean age children as distributed by age of the mother, this is considered to be the duration of exposure to mortality since the time they gave birth to the children. The second step is to calculate the proportion of women dead since they delivered which gives the proportion of the orphans. A backward projection of person-years lived by women is done by age and period. The third step is to take children ever born and calculate the distribution of surviving children belonging to a mother by their age and the age of the mother. Children ever born are then distributed by the same distribution. This steps aims at providing number of live births by year and age of the mother in single years. The fourth step aims at calculating ASFRs by age and period. Then agegroups and periods are merged as desired. Examples of the agegroups are 15-19, 20-24,....,45-49 and the recommended periods are 3 years, that is t-1 to t-3, t-4 to t-6 etc.

Since this method applies to a census/survey presented by year of age, all calculations are presented in discrete notations.

Let: w = age of women at census/survey (‘w’ from 12 to 64 years)

N(w) = Number of women age ‘w’ at census/survey

E(w) = Average number of children that women aged ‘w’ have ever given birth to at census/survey

C(w,k) = Total number of children aged ‘k’ and belonging to women age ‘w’ at census/survey, and

The distribution of children age ‘k’ to women age ‘w’=

$$C(w,k) = \frac{C(w,k)}{\sum xC(w,k)} \tag{3.1}$$

m(k) = Proportion of children age ‘k’ whose mother died (proportion orphan)

Calculations of fertility rates by age and period require the following steps:

(i) Estimate the relationship between proportion of surviving mothers and age of child. The estimation is done by a logit-linear relationship.

$$Logit(1–m(k)) = b0 + b1 * K \tag{3.2}$$

where b0 and b1 are obtained by linear regression,

(ii) Calculate the mean age of children living with mother, by age of mother: A(w). This helps in assessing mother's vulnerability to the risk of mortality from the birth to the census.

$$A(w) = \sum kK * \frac{C(w,k)}{\sum xC(w,k)} \tag{3.3}$$

(iii) Calculate the survival probability in the cohort of women now age 'w', from proportion of orphans of the same age A(w).

$$Logit(p(w)) = b0 + b1{*}A(w) \tag{3.4}$$

(iv) Backward project the number of person-years women have lived for the period preceding the census/survey, by age at census/survey, using the survival probability.

$$W(w-k, t-k) = \frac{N(w)}{p(w,k)} \tag{3.5}$$

(v) Backward project number of births as distributed by age of mother and time period, using the total children ever born and the distribution by age of child and age of mother. This assumes that surviving children living with their mother have the same age distribution as others.

$$B(w-k, t-k) = W(w) * E(w) * c(w,k) \tag{3.6}$$

(vi) Calculate the age specific fertility rates as distributed by age in single years and single year period.

$$F(w-k, t-k) = \frac{B(w-k, t-k)}{W(w-k, t-k)} \tag{3.7}$$

(vii) Merge age groups and periods as desired.

### 3.3.2 Limitations

#### 3.3.2.1 Age reporting

The major limitation in the application of own-children method for data in Sub Saharan Africa and majority of other developing countries has been reported as age misreporting (Cho et al., 1986); (Abbasi-Shavazi, 1997). This error may severely distort fertility that is based on age pattern as well as the trends in fertility estimates. This eventually causes the fertility estimates derived from own children method to have year to year fluctuations.

One way to ascertain the improvement in the age reporting is to look at the age pyramid plotted by single year. The other way of looking into the effects of age misreporting in fertility estimates derived from own children method is by examining fertility trends for single years. In order to reduce any distortions that could arise from age misreporting it is recommended to aggregate the estimates that have been derived from data for several years. One way to overcome unusual single year fluctuations is to calculate fertility rates for three periods of five years each and thereby diminishing the effect of age-misreporting . Moreover, five year age groups are used to calculate age specific fertility rates, this indicates that when data is aggregated then age-misreporting errors are minimized especially for applications where age data is inaccurate.

#### 3.3.2.2 Mismatching errors

Cho et al. (1986) and Abbasi-Shavazi (1997) notes that fertility estimates from own children method suffer from bias that arises from mismatching. These errors are divided into three; mismatch, allocations of unmatched children and failure to record the existence of some children because they are living in a geographic area other than that in which their mother lives; but it has been noted that errors from these sources are smaller than those from age misreporting.

According to Rindfuss (1977), there is one possibility where a mother and her biological child could be living in a three generation household, that is, with her (infants) maternal grandmother. It could be that the biological mother could be in school leaving the infant in mother's care. In such a scenario, it is possible for the infant to be captured as the grandmother's child instead of the mother's child.

The extent by which this method is biased is dependent on if there are younger children who may not be staying with their biological mothers. This situation may be different between societies besides among population subcategories. Although some children may have left their homes, in principle their number should be small.

#### 3.3.2.3 Under-reporting

Mismatch errors may also stem from under matching. some of the examples of under matching include Under-reporting of children ever born and those that are surviving. Undercounting gives less serious errors than age misreporting, but it can impact on the fertility estimates for the census year or the two years preceding the census (Cho et al., 1986); (Abbasi-Shavazi, 1997). In the process of calculating fertility rates, own-children method allows for adjustment of data where children and women are under-enumerated. If both children and women are under-enumerated then own-children fertility estimates are not affected.

#### 3.3.2.4 The presence of non-own children

The presence of children who do not belong to a woman, that is, children whose mother could not be identified in the household, can affect fertility estimates generated from own children method. The presence of non-own children in the data does not affect the level of fertility, as these children are distributed based on the proportion of children whose mother's age has been identified in the census files. However, the problem arises when non own children are mis-allocated to older mothers due to mismatching thereby leading to distortion of fertility patterns for older ages (Cho et al., 1986); (Abbasi-Shavazi, 1997). In general, the assumption made is that age of mother of unmatched children is distributed similarly to the age of mothers that were identified. The other problem in own children method could occur if ages for younger mother living with non own children is proportionately distributed similar to younger women with own children. This produces reapplication that is similar to older maternal ages consequently leading to overestimation of fertility for older woman and underestimating fertility for younger ages. Nevertheless, effects on total fertility are deemed to be negligible.

#### 3.3.2.5 Mortality assumptions

The process of measuring own children method can be affected if mothers are dead , as children whose biological mothers are dead will not be assigned to their own mothers hence will be omitted during the calculation of averages; if maternal mortality rates are low then the problem does not arise that much. According to Cho (1973), alternative model life tables are employed and depict a variation in life expectancy rates of about 10 years, reflecting two substantially different levels of morality. The results also showed that estimated TFR differed by less than 5% and the differences in ASFR were also minimal, suggesting that mortality estimation errors have low sensitivity. Abbasi-Shavazi et al. (2005) posits that maternal mortality has been on a downward trend, from around 237 per 100,000 live births in 1976 to around 37 in 1996 and it is estimated to be around 30 in recent years.

Thus, the impact of maternal mortality on the fertility estimates generated from own children method, if at all, is negligible.

### 3.3.3 Reverse Survival Method Application

Reverse survival method is applied during the estimation of fertility rates if data from surveys or census did not have questions that collected fertility directly. If population is assumed to be closed to migration, then individuals aged y years are held as the survivors of the births that happened y completed years antecedently. This means that the number of births that occurred y years ago can be calculated, as long as life table survival probabilities from birth to age y can be estimated, this is denoted by ,Ly/l0. If the population is reverse-survived to its birth year and divided by an estimated total population for that year then crude birth rate is computed. General Fertility Ratio, is computed by dividing the reverse-survived population by an estimate of women of childbearing age. To estimate TFR, we combine estimates of past births and women based on the age as computed from reverse survival models with estimates gotten from age pattern of fertility.

The methods requires number of births to be estimated for each year before the survey/census. The total number of individuals aged y years in a census or survey represents the survivors originating from births that happened within a year centered on the date y+0.5 years preceding the survey/census.

Algebraically,

$$B_{y+0.5} = \frac{N_y}{{}^cL_y} \tag{3.8}$$

with y being number of year before the census/survey, and ranges from 0 to 14; $N_x$ is total children aged y and reported in the census/survey; Cohort survival factor which measures survivorship is denoted by ${}^cL_y$. It relies on mortality at consecutive ages in successive years before a census/survey. Suitable age group estimates of mortality is usually available from data used in estimation of fertility. Some of the estimates are indirect from census data and include data for children ever born as well as those of surviving children, also direct estimates from birth histories analysis gotten from fertility survey data.

Data from either summary or full birth history can be used to estimate cohort survivor-ship. However, estimation for cohort survivorships can also be done using period mortality indicators through relational logit system of model life tables that are found in the Excel template that is provided with Timæus and Moultrie (2012).

To compute an estimate of total fertility, a numerator is required and should be an estimated number of births for all the years preceding a survey or census, an estimate of the mid-year number of women categorized in 5 year agegroups should be used as the denominator. The equation shown below can be used to estimated the mid-year number of women categorized by 5-year agegroup for each period T+5 preceding an enquiry from the number at T

$$ {}_5N^f_{y\,,T+5} = \frac{{}_5N^f_{y+5,T}}{{}_5P_{y,T}} \tag{3.9} $$

y denotes the 1st 5-year age group, and ranges from 10 to 60; T ranges from 0 to 10 and denotes time period; 5Py,T denotes survivorship at time T between the 5 year agegroups, and the following equation shows how it is computed

$$ {}_5P_{y,T} = \frac{{}_5L_{y+5}}{{}_5L_y} \tag{3.10} $$

With the availability of population estimates for periods between time period T prior a census or survey, then mid-year female population for each of the period can be estimated by linear interpolation.Total fertility estimates for each year prior to research is obtainable through distribution of total births for a given year for all the 5 year age cohorts of the same year. If total fertility equals one child per woman, then he equation shown below gives the expected number of births where the proportion of total fertility for each age group should be estimated with a view to compute the expected number of births to women in that agegroup.

$$_5{B^*}_{z,y+0.5} = {_5N}_{y+0.5} \cdot {_5f^*}_{z,y+0.5} \tag{3.11}$$

y denotes the number of years prior to the research and ranges from 0 to 14; z ranges from 15 to 45 and denotes the age groups, ; and percentage age specific fertility rates (PASFR) is denoted by this equation ${_5f}^*_{z,y+0.5}$ and is scaled to one.

The total number of expected births for all years prior the research a survey can be derived with ease.

$${B^*}_{y+0.5} = \sum_{z=15}^{45} {_5N}_{y+0.5} \cdot {_5f}^*_{z,y+0.5} \tag{3.12}$$

After obtaining the number of estimated births for each year prior to research, total fertility estimates can be obtained as shown in the following equation

$$TF_{y+0.5} = \frac{B_{y+0.5}}{B^*_{y+0.5}} \tag{3.13}$$

y denotes the number of years preceding the research and ranges from 0 to 14. Reverse survival mainly depends on the population distributed by age and sex, and as discussed here suffers from the poor quality data emanating from omissions, age heaping and other errors and thus the computations directly affects the resultant fertility estimates.

In addition, migration flows can affect the estimation of fertility rates by distorting either the numerator or the denominator or both. This study will utilize the Excel template FE_reverse_4.xlsx, provided with Timæus and Moultrie (2012), in order to appraise the suitability of the reverse survival in the estimation of fertility. The template is flexible since it has the capability of selecting different mortality models with a purpose to reverse-survive the population. Researchers have the option of choosing one of the four Coale-Demeny Model Life Tables, or utilize empirical mortality pattern or use the United Nations General Pattern. Moreover, for mortality changes that happened prior the research be to reproduced, it is possible to modulate the level and age pattern of mortality in the three quinquennial periods prior to the research by entering Spoorenberg: Reverse survival method of fertility estimation: An evaluation values of under-5 $({}_5q_0)$ and adult mortality $({}_{45}q_{15})$ or of alpha and beta parameters.

### 3.4 Data Quality

When carrying out research social scientists face many questions concerning the use of a particular data set, key among them is; does collected data suffer from errors and what could be the nature and magnitude, with specific reference being made to the main variables that matter most in the study? This question plays a vital role in this research, since the quality of collected data determines if the fertility estimates from such data will be useful and reliable.

This study used graphs, the age ratio method, Whipples and the Myer's blended method and also the distribution of children ever born to evaluate age data. Data collected from surveys and census suffer from age errors as a result of coverage errors, which could emanate from failure by the enumerator to record the correct age as well as errors in age misreporting arising from the respondent. In particular, age data from sub saharan Africa and other developing countries is found to have high preference for some digits as a result of ignorance and illiteracy (Hlabana, 2010). In order to test age accuracy for data in single years, Myers' blended index is mostly used because it aids in identifying if ages ending in some digits are preferred over others. The results from Myer's index helps in identifying if the ten digits between 0 and 9 both being included are either preferred or avoided.

The population is blended in a way that each digit has similar summation, that is , they have equal sum, with each of the ten digits having a blended total that is almost 10% of the overall total. Hence, the computed index for a given digit indicates the deviation from 10% of the proportion for entire population; the index shows if a particular digit is over-selected or avoided. If there is no age heaping, then the index will be 0, but if the index will be 90, it indicates that all ages have been reported at a single digit.

Shryock and Siegel (1976) on the other hand, noted that in order to check the extent of errors in age reporting, then age-groups are used in the application of age ratio method. Age ratio method is denoted as the ratio of the population for a certain age group to 1/3 of the populations summation in the same age group and the preceding and next groups, multiplied by a hundred. The age index computed by this method makes assumptions that any variations from a hundred suggest a shift in reporting age; an index of below 100 imply that age has been understated while an index above 100 imply that age has been overstated.

While it is recommended that data should be evaluated and assessed soon after data processing, evaluation and assessment of data is a recursive activity at all stages during analysis, data users are cautioned to utilize tabulated results with a skeptical eye, and being keen to some errors and bias that might be introduced by the processed data into the final results.

Census and survey data usually suffer content errors that arise from age and parity misreporting which can be as a result of age heaping or age exaggeration (Moultrie and Zaba, 2013). For this study, quality of data was assessed using graphs, calculating the indices of both

Whipples and Myers blended as well as assessing the age ratio and lastly by use of distribution of children ever born to a woman.

### 3.4.1 Age Misstatements

Graphing of the population data as distributed by age in single years and by sex is beneficial to a researcher since any age heaping prevalent in the data is made visible at an early stage. Age heaping can be assessed visually and is deemed as a good indication of ages being piled together or heaped and is comparable to the assessment made from measures such as Whipple's Index, Myers' Blended Index and United Nations Age-Sex Accuracy Index. A manual from US Census Bureau concludes that while Whipple's Index, Myers' Blended Index and United Nations Age-Sex Accuracy Index are helpful as they provide summary statistics or useful during comparisons they fail to give an understanding into flaws that could be present in the data which are not obtainable through ratio and graphical analyses of the data. According to the US Census Bureau(1985) age heaping is exhibited by concentrating ages ending in 0 and 5 years for a given age distribution. Age heaping could equally occur for other ages based on how census data is collected or derived.

When demographic change is systematic due to absence of large exogenous events, the number of people found during the enumeration for all ages is expected to have a smooth progression. It is notable that fertility has always been high in developing countries and therefore population size is expected to decrease in a monotonic way by age. When number of births in absolute figures are found to have declined in recent years, then it is expected that children at younger ages would be slightly lower than at older ages.

One of the disadvantages of graphing population as distributed by age and sex is that any misrepresentation and error in the data for older ages tend to be concealed by those in younger ages due to their larger population sizes. The alternative methods for exploring the misrepresentations and errors in data for older ages is ratios or relative rates. In the absence comparative data, it is recommended that higher age ranges should be considered on an individual basis.

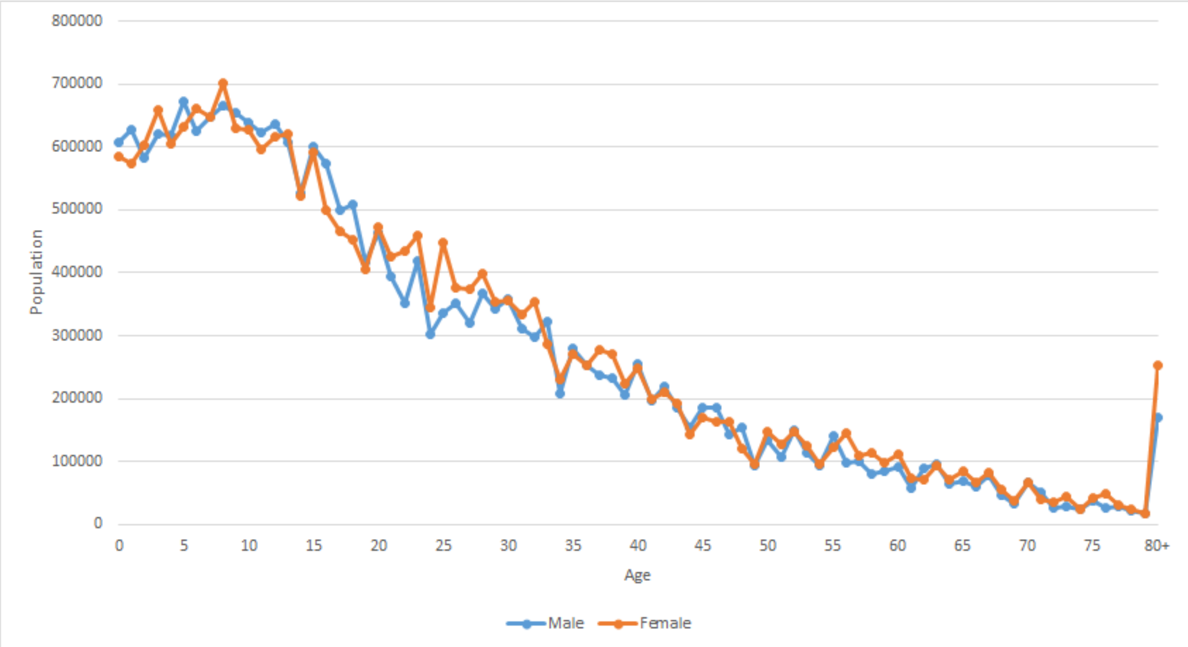
Age misstatement end up misplacing births and parities and hence distorts fertility estimation when using different methods. During estimation of fertility, it is therefore crucial to lessen the effect of the errors. During the 2015/16 KIHBS survey, women in the reproductive age groups were asked to report both their age with both month and year of birth questions being asked to act as a crosscheck. According to Shryock and Siegel (1976), the UN recommends that census and surveys should use either date of birth or age at last birthday when collecting data on age. It is noted that errors in age reporting are still persistent in both surveys and censuses.

According to Kpedekpo (1982) women in older ages are inclined to exaggerate their ages by overstating them. In many African societies, the elderly tend to be treated with high respect. This acts as a motivation for some older women to be inclined to exaggerate their age which comes with some associated respect.

As a result, by exaggerating of the age of mothers this leads to some births being reported for women who are already above the reproductive ages which is mostly beyond 50 years and if fertility is estimated using conventional manner then they are consequently excluded (Zaba, 1981). In addition, this leads to fertility rates inflation reported by women beyond 40 years.

As earlier pointed out,graphing is another visual way to identify if the data has age misreporting. Figure 3.1 shows population distribution in single years for both sexes. The graph shows that there was no strong preference for digit 0 and 5 as compared to the other digits as the determined peaks and troughs were also present in those other digits. This indicates that there was no significant age misreporting or heaping in the data.

Figure 3.1: Age distribution in single years




Source:Analysis of 2015/16 KIHBS

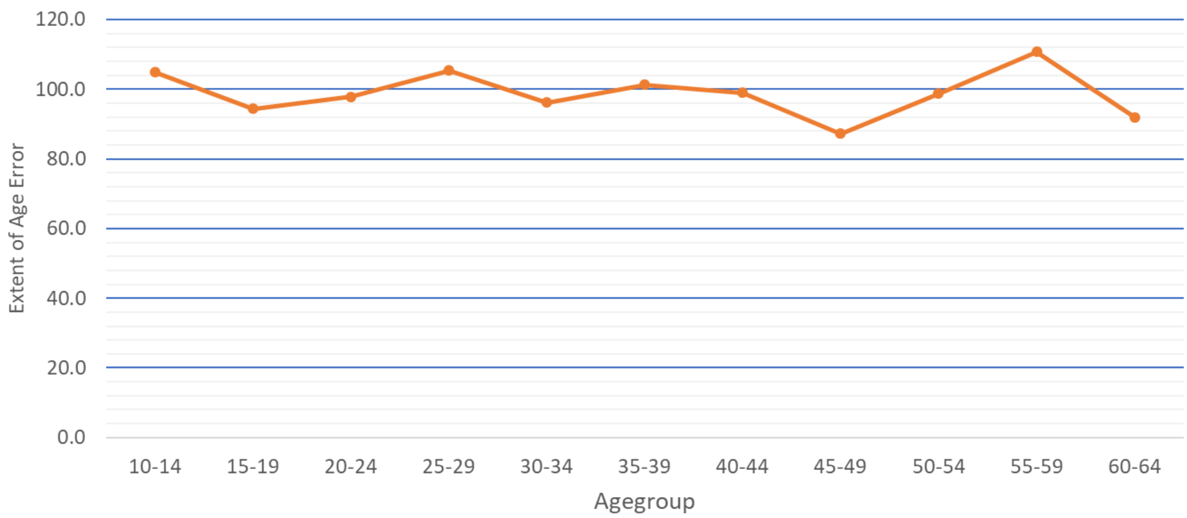
### 3.4.2 Age ratios

Age heaping in some ages is easy to identify through graphs as compared to measures that need calculations. However, age ratio calculations tend to give an essential indication of probable undercounts or age displacements. Age ratio method is denoted as the ratio of the population for a certain age group to 1/3 of the populations summation in the same age group and the preceding and next groups, multiplied by a hundred. The age index computed by this method makes assumptions that any variations from a hundred suggest a shift in reporting age; an index of below 100 imply that age has been understated while an index above 100 imply that age has been overstated.

Given that the population change is expected to be fairly linear between age groups, then age ratio is expected be reasonably close to 100. If there no plausible exogenous factors for example, past disasters or tragedies and migration that could affect particular age groups then any deviations from 100 then it indicates that the data could have undercount or displacement errors. Any distortions in the data for any particular age-group is likely to cause disruptions in the age ratios in either side of the age group. When an age-group has low figures, then age ratio for that group will be below one making the adjacent age-groups to have spikes.

Figure 3.2 indicates some minimal erratic overstating and understating of age for almost all the reproductive-aged women interviewed during 2015/16 KIHBS. This erratic behaviour for the younger and older women refutes the allegations that the data could have suffered from age understatement and overstatement. If age understatement or overstatement exists then it is expected to be seen over 100 for older and under 100 for younger women. The most noticeable age errors were seen for women in the 45-49 agegroup and 55-59 agegroup, with age reporting error being about 11% . The errors are assumed not to be systematic and therefore their effects on fertility estimates would cancel out.

Figure 3.2: Extent of Age Error detection by agegroup




Source:Analysis of 2015/16 KIHBS

### 3.4.3 Age Heaping

Kpedekpo (1982) posits that some people have the tendency of showing preference in reporting certain ages over others. This tendency is called heaping or digit preference. While it is expected that digit preference can occur at any digit, but evidence shows that 0 and 5 digits are highly preferred. In order for avoidance or prevalence of each terminal digit to be detected, the Whipples index is employed. The generated index ranges between 100 and 500. Collected data is deemed highly inaccurate if index is more than 175, when it varies from 125 to 175, it is said to be inaccurate, if it ranges between 110 and 125 then it is treated as fairly acceptable and if below 105 it is deemed to be highly accurate. The value of the index calculated from the 2015/16 KIHBS data was 49.0 and 57.5 for terminal digit 0 and 5, respectively. This implies that the data was highly accurate and did not require adjustment before use.

However, Myer's blended method is recommended in order to detect preference for all terminal digits. This method generates an index of preference for digit 0 to 9 which represents the deviation from 10% of the proportion of the total population that reports for each digit.

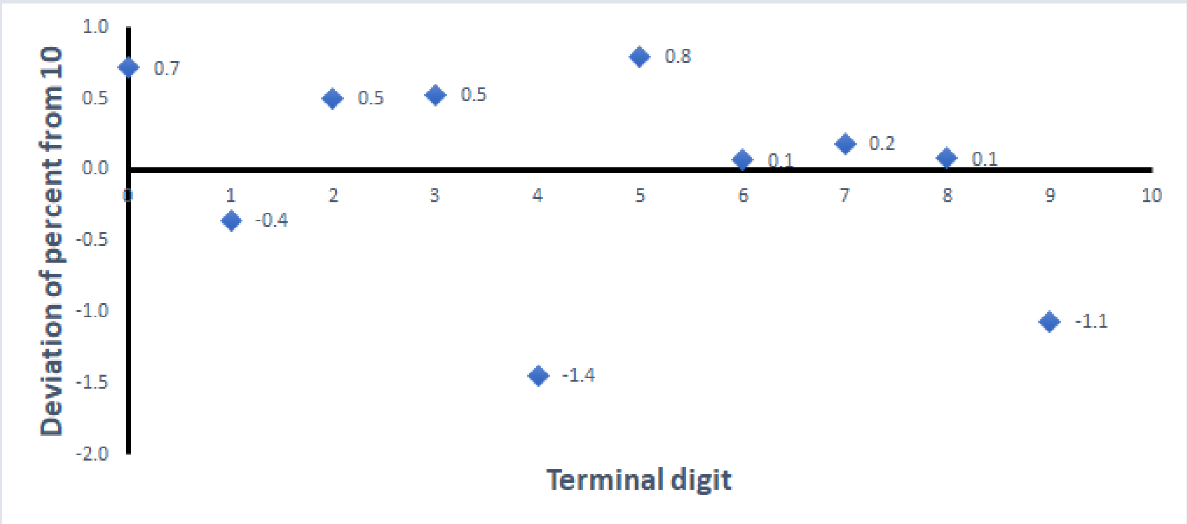
From Table 3.1 and Figure 3.3 below, Myer's blended method was calculated to be 2.9. The data shows insignificant preference and avoidance of ages for different terminal digits.

Table 3.1: Results from Myer's Blended

| Terminal digit | Blended Population | Percent distribution | Deviations from 10 |
|---|---|---|---|
| 0 | 14,571,752 | 10.7 | 0.7 |
| 1 | 13,108,798 | 9.6 | -0.4 |
| 2 | 14,283,194 | 10.5 | 0.5 |
| 3 | 14,319,925 | 10.5 | 0.5 |
| 4 | 11,632,192 | 8.6 | -1.4 |
| 5 | 14,680,318 | 10.8 | 0.8 |
| 6 | 13,690,160 | 10.1 | 0.1 |
| 7 | 13,850,001 | 10.2 | 0.2 |
| 8 | 13,705,569 | 10.1 | 0.1 |
| 9 | 12,150,383 | 8.9 | -1.1 |
| **Total** | **135,992,291** | **100.0** | **5.7** |

Source:Analysis of 2015/16 KIHBS

Figure 3.3: Myers index for digit preference,2015-2016

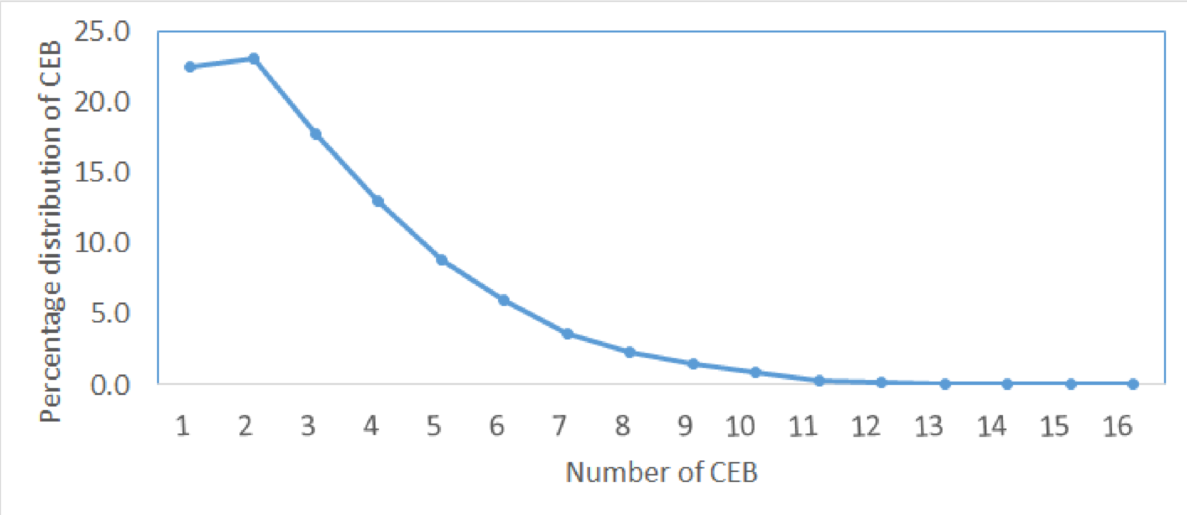



Source:Analysis of 2015/16 KIHBS

Kpedekpo (1982) notes that in Africa, error in age reporting is common and this often happen if the research assistant skips the question or if the respondent does not know their age. Preston (2001) also notes that even after interviewers trying to probe in order to solicit the information, some of the respondents fail to provide the information. If data is unavailable or missing, then demographic analysis faces some challenges. Missing data has been theorized to include cases that are deleted, likelihood-based estimation, imputation for non-response. However, Little and Rubin (1987) have recently theorized that there is a shift from treating missing data as a thing to be deleted but rather should be included in the analysis to help examine any response biases. They noted that if there is some correlation between missing cases the outcomes of interest, leaving missing cases would introduce bias in the analysis.

Based on 2015/16 KIHBS, 22,978,493 females were covered in the survey, and only 21,802 (0.1%) women did not give information on their age. There were 12,778,472 female respondents aged 12-49 years and (58.4%) of the women indicated that they had ever given birth to 24,264,386 children. This study assumed that there was no significant missing data in the age variable and hence there was no imputation done in the analysis. Figure 3.4 depicts the distribution of Children Ever Born (CEB) by number of CEB. The figure shows that majority of the women reported to have ever given birth to 3 or less children, and the number of women decreases when the number of children increases.

Figure 3.4: Distribution of Childen Ever Born




Source:Analysis of 2015/16 KIHBS

### 3.4.4 Lifetime Fertility

In order to estimate lifetime fertility, one of the prerequisite is details on the aggregate number of children a woman has ever given birth to. According to Muhwava and Timæus (1996), some factors such as age misreporting, omission of children who are dead or those not living with their mother and inclusion of still-births can affect information on children ever born as shown in Table 3.2. Past research has shown that the direction of bias for both emigrants and immigrants is dependent on if the fertility for both categories is considerably variant from non-migrants fertility rates. There is always some probability of missing these women during data collection since they are always mobile or due to failure to have representative proxies.

Table 3.2: Lifetime Fertility errors and Direction of the Bias

| Type of error | Direction of Bias |
|---|---|
| 1. Dead women omitted | + |
| 2. Emigrant women omitted | + |
| 3. Immigrant women omitted | + |
| 4. Dead or absent children omitted | - |
| 5. Adopted or step children included | + |
| 6. Zero parity reported as not stated | + |
| 7. Not stated reported as zero parity | - |
| **Net Effect** | **Negative (increasing with age)** |

Source:Muhwava and Timæus (1996)

In the 2015/16 KIHBS, 20,189 women in the 12-49 age bracket failed to report their parity. The failure to report parity might be as a result of errors during data entry for women with zero parity and is more prevalent to young women (United Nations, 1983). The age distribution of these women with zero parities did not show any bias and was represented across all the ages in the reproductive group. The proportion of women with zero parity was insignificant and were therefore not included in the analysis. From the analysis, parities varied from zero to 16 children. While it is expected that variation in fertility exists among different women, in 2015/16 KIHBS, some women reported CEB which was biologically impracticable as shown in Table 3.3 below. For example, there are 994 women aged 17 who reported to have had 12 children. There are other similar unrealistic cases that were reported and were therefore not used in the estimation of fertility since they would distort the estimates.

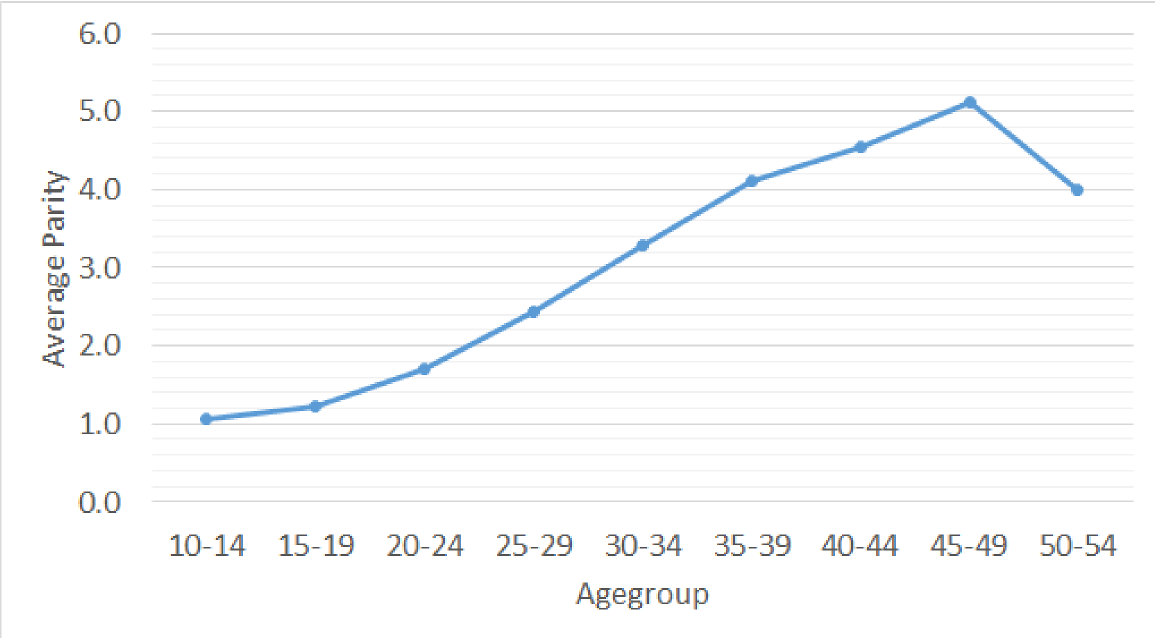
Table 3.3: Children Ever Born by Age of Mother

| | Children born alive | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Age | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| 12 | 1,881 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 13 | - | 520 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 14 | - | 2,775 | 877 | - | - | 348 | - | - | - | - | - | - | - | - | - | - | - |
| 15 | 1,977 | 9,816 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 16 | 53 | 15,130 | 631 | 110 | - | 83 | - | - | - | - | - | - | - | - | - | - | - |
| 17 | - | 41,571 | 4,252 | 895 | 66 | - | - | - | - | - | - | - | 994 | - | - | - | - |
| 18 | 1,005 | 61,738 | 12,872 | 1,108 | 512 | - | - | - | - | - | - | - | - | - | - | - | - |
| 19 | - | 87,564 | 13,832 | 2,573 | 881 | - | - | - | - | - | 469 | - | - | - | - | - | - |
| 20 | - | 115,421 | 40,568 | 9,403 | 1,416 | 493 | 30 | 387 | - | - | - | - | - | - | - | - | - |
| 21 | - | 117,162 | 60,591 | 14,212 | 4,239 | 2,410 | 596 | - | - | - | - | - | - | - | - | - | - |
| 22 | - | 131,894 | 73,437 | 22,862 | 6,836 | 2,741 | - | - | - | - | - | - | - | - | - | - | - |
| 23 | 1,359 | 125,506 | 115,171 | 39,082 | 11,931 | 5,392 | - | - | - | - | - | - | - | - | - | - | - |
| 24 | 434 | 104,770 | 71,622 | 35,032 | 16,190 | 7,415 | 1,678 | - | - | - | - | - | - | - | - | - | - |
| 25 | 1,575 | 140,546 | 129,061 | 59,823 | 24,556 | 8,991 | 4,159 | 470 | - | - | - | - | - | - | - | - | - |
| 26 | 1,557 | 101,938 | 105,479 | 51,441 | 24,969 | 11,794 | 8,089 | 1,347 | 1,213 | - | - | - | - | - | - | - | - |
| 27 | 164 | 100,031 | 108,229 | 54,717 | 34,943 | 15,909 | 7,997 | 4,552 | 653 | 120 | - | - | - | - | - | - | - |
| 28 | 2,377 | 79,694 | 105,304 | 69,516 | 48,982 | 25,513 | 16,446 | 4,948 | 2,008 | 709 | - | - | - | - | - | - | - |
| 29 | 152 | 64,630 | 106,144 | 70,310 | 34,547 | 29,333 | 8,991 | 7,262 | 1,887 | 947 | - | - | - | - | - | - | - |
| 30 | 611 | 40,498 | 73,668 | 76,163 | 71,369 | 33,869 | 18,412 | 7,962 | 6,829 | 1,101 | 633 | - | - | - | - | - | - |
| 31 | - | 36,727 | 97,626 | 60,781 | 43,893 | 33,272 | 17,252 | 4,434 | 650 | 1,837 | 630 | - | - | - | - | - | - |
| 32 | 1,176 | 49,607 | 97,243 | 68,498 | 56,847 | 30,472 | 17,380 | 7,722 | 4,621 | 1,360 | 448 | 277 | - | - | - | - | - |
| 33 | - | 35,227 | 61,140 | 59,812 | 46,829 | 37,291 | 19,407 | 10,294 | 5,387 | 2,081 | 1,400 | - | 187 | - | - | - | - |
| 34 | 1,982 | 21,474 | 45,768 | 66,842 | 33,943 | 20,565 | 18,253 | 5,710 | 3,291 | 949 | 55 | 550 | - | - | - | - | - |
| 35 | 380 | 16,279 | 57,168 | 54,817 | 51,248 | 34,026 | 22,936 | 14,376 | 13,140 | 1,463 | 2,026 | 412 | - | - | - | - | - |
| 36 | - | 34,763 | 37,609 | 57,027 | 39,970 | 28,577 | 15,244 | 13,472 | 6,412 | 5,418 | 1,519 | - | 184 | - | - | - | 204 |
| 37 | - | 18,641 | 47,215 | 54,167 | 49,239 | 41,644 | 21,329 | 15,737 | 11,293 | 2,658 | 3,506 | 2,358 | 505 | - | - | - | - |
| 38 | - | 10,837 | 29,194 | 55,847 | 53,865 | 34,905 | 34,573 | 17,486 | 7,487 | 6,415 | 3,683 | 341 | 820 | - | - | 15 | - |
| 39 | - | 7,831 | 24,764 | 37,204 | 55,960 | 33,967 | 24,253 | 18,558 | 9,699 | 5,247 | 1,246 | 544 | - | - | - | - | - |
| 40 | 584 | 16,048 | 36,952 | 29,667 | 47,129 | 26,230 | 29,504 | 22,285 | 8,451 | 13,155 | 5,455 | 3,428 | 460 | 232 | - | - | - |
| 41 | - | 14,944 | 27,470 | 37,841 | 26,704 | 23,570 | 18,367 | 15,722 | 13,178 | 7,017 | 2,158 | 1,279 | 1,179 | - | 1,064 | 369 | - |
| 42 | 2,229 | 22,494 | 17,432 | 46,022 | 29,456 | 25,664 | 23,442 | 16,505 | 8,365 | 6,287 | 4,012 | 1,922 | 682 | 539 | - | - | - |
| 43 | 337 | 13,537 | 36,652 | 33,749 | 25,496 | 21,919 | 20,782 | 9,153 | 12,369 | 6,014 | 4,195 | 3,230 | 633 | - | - | - | - |
| 44 | - | 5,968 | 9,172 | 35,836 | 25,062 | 18,279 | 15,588 | 9,737 | 8,271 | 5,607 | 3,228 | 1,120 | 1,701 | 382 | - | - | - |
| 45 | 116 | 7,296 | 15,502 | 30,739 | 21,459 | 26,672 | 17,238 | 11,621 | 11,356 | 12,034 | 7,070 | 2,005 | 3,212 | 730 | - | - | - |
| 46 | 122 | 6,224 | 20,773 | 23,860 | 20,456 | 19,667 | 26,175 | 13,139 | 13,840 | 4,831 | 4,654 | 1,408 | 452 | 980 | 240 | - | - |
| 47 | 121 | 6,904 | 14,139 | 16,994 | 28,446 | 29,807 | 16,994 | 12,672 | 10,929 | 8,458 | 7,856 | 1,742 | 1,327 | - | - | - | - |
| 48 | - | 5,540 | 10,956 | 26,969 | 21,114 | 14,487 | 11,627 | 10,066 | 7,521 | 5,246 | 3,311 | 472 | 589 | 157 | - | - | - |
| 49 | - | 1,208 | 9,951 | 15,518 | 9,657 | 16,211 | 10,267 | 13,083 | 4,937 | 8,449 | 4,316 | 1,584 | 178 | - | - | - | - |

Source:Analysis of 2015/16 KIHBS

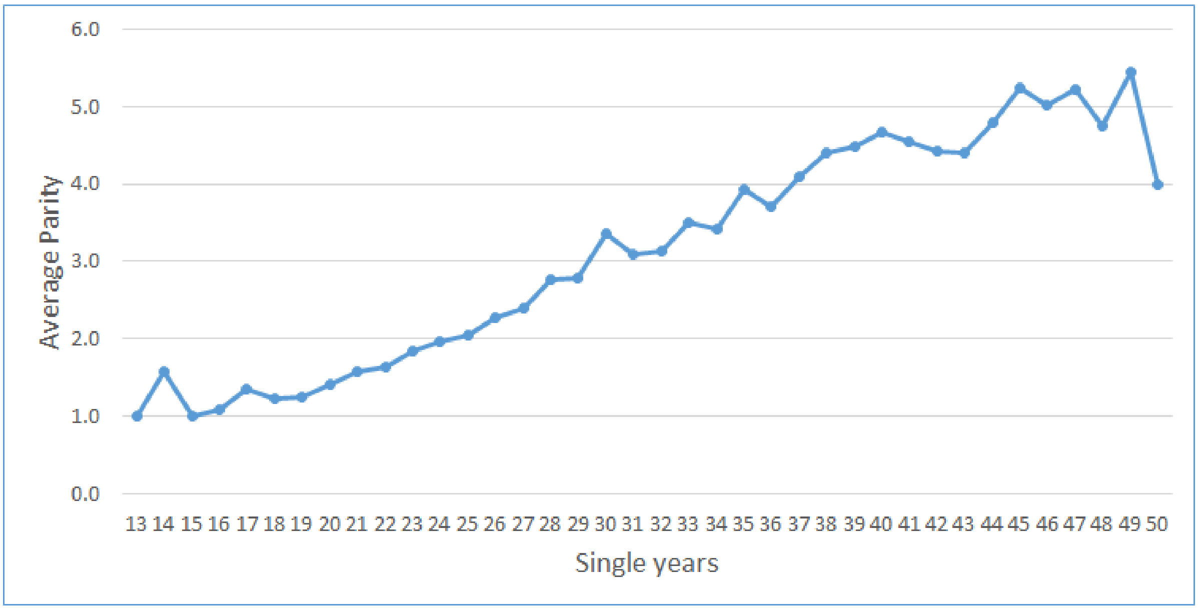
Figures 3.5 and 3.6 show average parity as plotted by single ages and also by age groups, respectively. For both graphs, the average parity curves indicate a uniform increase in children ever born by age. Brass (1981), posits that average parity distribution should normally increase with age, and diminish towards the end of reproductive period. The average parity pattern was found to be acceptable for analysis. The average parities curve for CEB by single ages tends to meander from age 30 years as we tend to old ages. Unless when fertility is affected by irregular changes, the meandering indicates that as young women are consistent in the reporting of their parities as compared to older women. A curve depicting average parities by age-groups shows that there is uniform rate of increase in CEB as women move into higher age-groups. This figures show that when there are no changes in levels of fertility, average parities tend to increase with age since more women could be going back to childbearing. However the rate of increase tend to diminish as the women near the end of reproductive period since majority of them will be nearing the end of childbearing.

Figure 3.5: Average Parity distributions by Agegroup




Source:Analysis of 2015/16 KIHBS

Figure 3.6: Average Parity Distribution by Single years




Source:Analysis of 2015/16 KIHBS

2015/16 KIHBS shows that in total, 24,264,386 children were reported to have ever been born. However, when the women were asked detailed information on the number of children still living with them, number of children living elsewhere and and number of children who are dead by their sex orientation, a total of 24,339,839 children was reported as depicted in Table 3.4. The difference in reporting was 0.31 per cent makes the reporting of parities in the 2015/16 dataset satisfactory.

Table 3.4: Detailed Information on Total Number of Children Ever Born

| Variable | Cases |
|---|---|
| Ever Given Birth | 7,464,876 |
| Never Given Birth | 5,313,596 |
| Children Ever Born | 24,264,386 |
| Sons Ever Born | 12,221,119 |
| Daughters Ever Born | 12,043,267 |
| Sons Living with Mothers | 9,824,729 |
| Sons Living Elsewhere | 1,746,142 |
| Sons Dead | 720,273 |
| Daughters Living with Mothers | 9,398,967 |
| Daughters Living Elsewhere | 2,009,160 |
| Daughters Dead | 640,569 |
| Total Children | 24,339,839 |

Source:Analysis of 2015/16 KIHBS

In summary, 2015/16 KIHBS data reported a Whipples index of 49.0 and 57.5 for terminal digits 0 and 5 respectively. Myer's blended index was 2.9 and this was an indication that in general the data was accurate and therefore did not require any adjustment to improve its quality before use. Table 3.3 also shows some 17 year old women who have given birth to more than 7 children which makes partial birth history produce inconsistent results. Failure to exclude such outliers during the estimation of fertility could lead to inaccurate results.

# Chapter 4

# Levels and trends in Fertility in Kenya

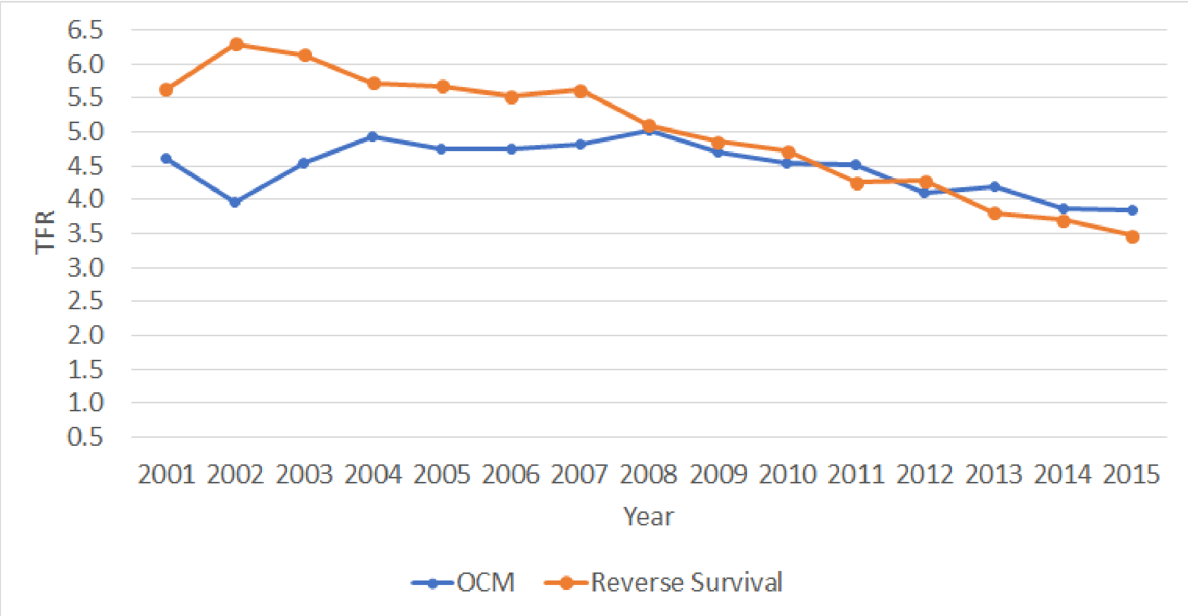
## 4.1 Introduction

This chapter presents the reconstructed trends in fertility based on 4-parameter OCM and reverse survival methods. The results were then compared to those obtained from birth history.

### 4.1.1 National estimates of total fertility rates (TFR) by own children method (OCM) and reverse survival

Figure 4.1 below shows the estimated levels and reconstructed trends of TFRs between 2001 and 2015. The results show that fertility trends using the reverse survival techniques were systematically higher than those generated from the OCM models for years between 2001 and 2010. The two methods showed a uniform trends for the period between 2010 and 2015 but reverse survival technique generated lower TFRs as compared to OCM. This shows that the two techniques are capable of producing consistent trends when the years are closer to the survey date as compared to when you move further from the survey date.

Figure 4.1: Reconstructed trends and estimates of TFRs




Source:Analysis of 2015/16 KIHBS

Table 4.1 shows TFR and GFR results of the reverse survival and OCM methods when applied to 2015/16 KIHBS dataset. The results show both techniques generated TFR estimates that ranged from 3.5 to 6.3 and 3.8 to 5.0 for reverse survival and OCM, respectively, for the 2002 to 2016 period. In majority of the years, OCM had TFR estimates that were systematic lower than reverse survival estimates. The fertility trend indicates that Kenya had a fertility stall in 2002 which supports the literature that Kenya experienced a fertility stall between 1993 to 2003 (Westoff et al., 2006) and (Mutuku, 2013). In the period under review, the average number of children being born to women aged 15 - 49 years was consistently low for OCM estimates.

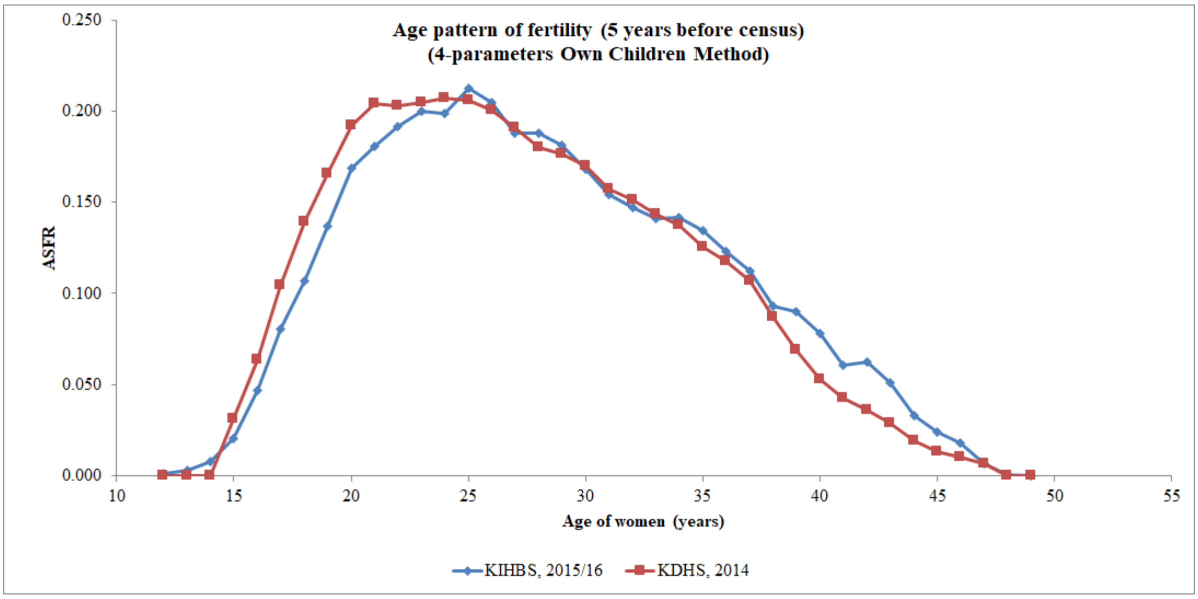
Table 4.1: Fertility estimates per year(2002-2016)

| | Reverse Survival | | OCM | |
|---|---|---|---|---|
| Mid-Year | TFR | GFR (15-49) | TFR | GFR (15-49) |
| 2016 | 3.5 | 114 | 3.8 | 123 |
| 2015 | 3.7 | 122 | 3.9 | 126 |
| 2014 | 3.8 | 126 | 3.8 | 124 |
| 2013 | 4.3 | 141 | 4.2 | 134 |
| 2012 | 4.3 | 141 | 4.1 | 132 |
| 2011 | 4.7 | 156 | 4.5 | 142 |
| 2010 | 4.9 | 160 | 4.5 | 141 |
| 2009 | 5.1 | 167 | 4.7 | 148 |
| 2008 | 5.6 | 183 | 5.0 | 158 |
| 2007 | 5.5 | 179 | 4.8 | 152 |
| 2006 | 5.7 | 184 | 4.7 | 149 |
| 2005 | 5.7 | 185 | 4.7 | 152 |
| 2004 | 6.1 | 198 | 4.8 | 157 |
| 2003 | 6.3 | 203 | 4.4 | 145 |
| 2002 | 5.6 | 181 | 3.9 | 130 |

Source:Analysis of 2015/16 KIHBS

Figure 4.2 presents a comparison for 5 year averages of ASFRs using OCM technique, for both 2015/16 KIHBS and 2014 KDHS. The ASFRs obtained from the 2015/16 KIHBS and 2014 KDHS show a uniform trend for ages 12 to 49 years. KIHBS data produced ASFRs that were lower for women aged 15 to around 25 years. The two dataset reported similar ASFRs for the women aged betwen 25 and 35 years. Figure 4.2 also shows that the trends changes for women aged 35 to 49 years where 2014 KDHS dataset reports slightly lower ASFRS. The figure shows that OCM produces schedules of ASFRs that are stable based on the smoother graphs. The ability of own children method to produce smoother schedules of ASFRs gives it some strength that can be exploited at lower levels of aggregation to produce localised fertility estimates. The consistent schedules of ASFRs from Figure 4.2 imply that fertility levels in Kenya obtained from 2015/16 dataset are consistent and therefore reliable.

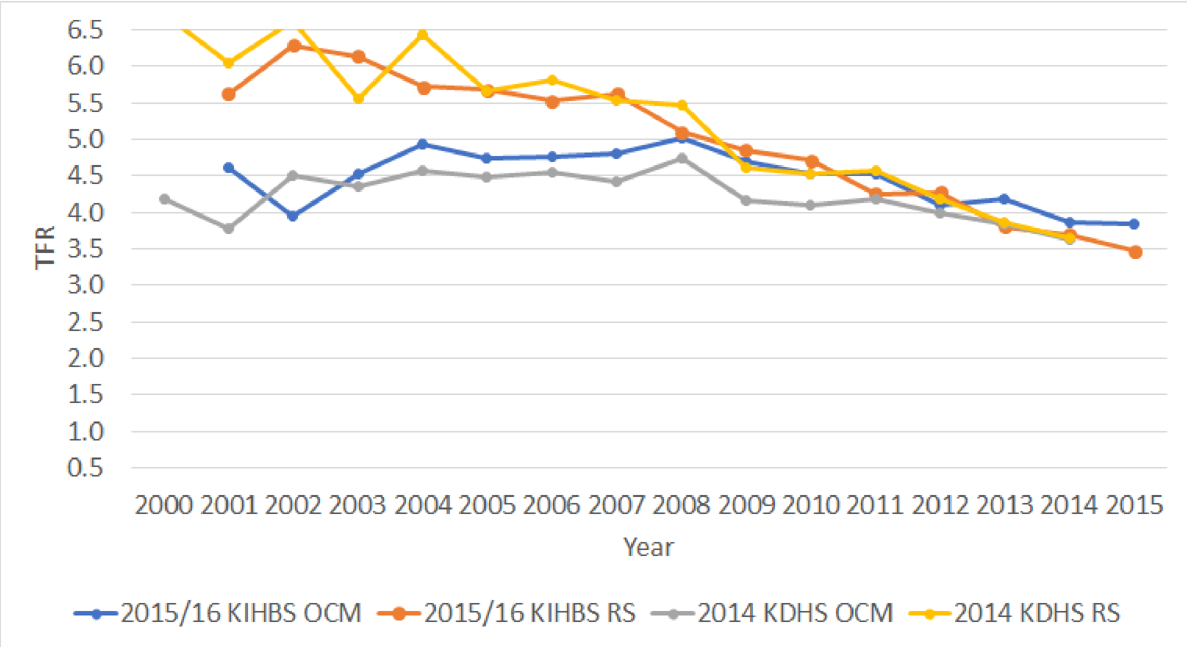
Figure 4.2: Comparison for 5 year Averages of ASFRs using OCM technique




Source:Analysis of 2015/16 KIHBS

The comparison of TFRs obtained from both 2015/16 KIHBS and 2014 KDHS datasets using own children method and reverse survival method is presented in Figure 4.3. The estimated levels and reconstructed trends of TFRs indicate that the two methods produces a similar and uniform trend for the five years preceding both surveys period. For the remaining TFRs estimated using the reverse survival method were consistently higher than those obtained from the own children technique which is expected. Figure 4.3 shows that in the fourth and fifth years preceding the survey, the gap between TFR estimates from the two methods appears to widen.

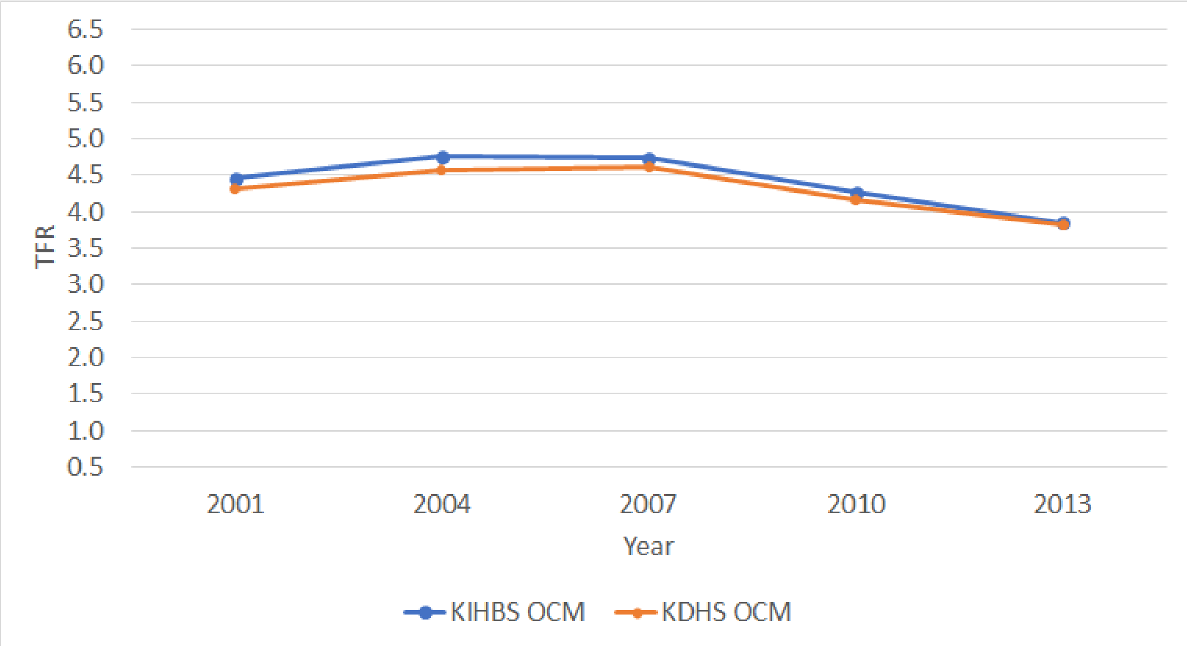
Figure 4.3: Comparison TFRs obtained from 2015/16 KIHBS and 2014 KDHS




Source:Analysis of 2015/16 KIHBS and 2014 KDHS

Figure 4.4 depicts a comparison of TFRS obtained from 5 year groups using own children method for both 2015/16 KIHBS and 2014 KDHS dataset. The TFRs obtained display a uniform trend for all the years preceding survey years, that is from 2001 to 2013. This indicates that the results obtained from own children method are consistent despite using different datasets which gives an indication that the estimated TFRs are reliable.

Figure 4.4: 5 Year groups Own Children Method




Source:Analysis of 2015/16 KIHBS and 2014 KDHS

Table 4.2 shows the estimated results of the reverse survival and OCM method when applied to 2014 KDHS dataset. The results indicate that TFR estimates range from 3.6 to 6.7. Both reverse survival and own children method reported TFR of 3.6 in 2014 as compared to birth history method that was reported at 3.9. The results supported what Avery et al. (2013) had posited that FHB produces significantly larger TFR as compared to OCM technique. The results also showed that as you move away from the survey period, reverse survival produces reconstructed TFR estimates that are systematically higher than OCM. The estimated TFR from 2014 KDHS were similar to those from 2015/16 KIHBS indicating that the estimates from the latter are reliable.

Table 4.2: Fertility estimates per year(2000-2014)

| | Reverse Survival | | OCM | | BH Method |
|---|---|---|---|---|---|
| Mid-Year | TFR | GFR (15-49) | TFR | GFR (15-49) | TFR |
| 2014 | 3.6 | 120.5 | 3.6 | 121 | 3.9 |
| 2013 | 3.9 | 127.9 | 3.9 | 133 | |
| 2012 | 4.2 | 138.0 | 4.0 | 137 | |
| 2011 | 4.6 | 151.0 | 4.2 | 148 | |
| 2010 | 4.5 | 149.2 | 4.1 | 146 | |
| 2009 | 4.6 | 151.2 | 4.2 | 152 | 4.6 |
| 2008 | 5.5 | 178.6 | 4.7 | 175 | |
| 2007 | 5.5 | 179.7 | 4.4 | 165 | |
| 2006 | 5.8 | 187.0 | 4.6 | 174 | |
| 2005 | 5.7 | 180.9 | 4.5 | 173 | |
| 2004 | 6.4 | 204.8 | 4.6 | 183 | |
| 2003 | 5.6 | 177.2 | 4.4 | 179 | 4.9 |
| 2002 | 6.6 | 211.1 | 4.5 | 192 | |
| 2001 | 6.1 | 193.0 | 3.8 | 169 | |
| 2000 | 6.7 | 214.1 | 4.2 | 196 | |

Source:Analysis of 2014 KDHS

## 4.2 Discussion

The purpose of this study was to determine the extent to which methods for estimating trends in fertility without use of birth history could be used on Kenyan surveys data. Results from 2015/16 KIHBS showed that reverse survival estimated TFR to be 3.5 as compared to OCM that estimated it to be 3.8. To check for the consistency of the results, 2014 KDHS was used as shown in Table 4.2. The results from 2014 KDHS dataset are consistent when using both reverse survival and own children method but under-reports TFR compared to birth history method. Both Table 4.1 and 4.2 indicate that even though the two models have different assumptions, derived estimates are within range.

The study showed that both techniques can reveal a trend in fertility for upto a period of 15 years. The results show that reverse survival can produce better current estimates as compared to own children method though both methods have an added advantage since they can reveal a trend in fertility.

The results from this study agree with what Spoorenberg (2014) had noted that reverse survival gets estimates that are robust and closer to what other indirect methods get. The results also confirm the notion that OCM has advantages over other indirect methods due to its non requirement of prior assumptions on past vital rates and also does not require any direct and detailed information on birth statistics. This study was an extension of the study done by Collins and Levin (2008). They had used the original OCM, while this study was based on the 4 parameter model developed by Garenne and McCaa(2017).

# Chapter 5

## Summary, Conclusion and Recommendations

### 5.1 Introduction

This chapter presents the summaries derived from the research findings, the conclusions and recommendations made based on the results.

### 5.2 Summary

The research aimed at determining fertility levels and trends in Kenya without use of birth history data by first appraising the quality of data . According to Brass (1996), credibility of estimated rates of fertility depends on the quality of collected data. For this study, 2015/16 KHIBS data was evaluated in order to check if the data met the threshold of being accurate in estimation of fertility rates.

When the data was examined using age ratios, it was noted that there was some minimal erratic overstating and understating of age for almost all women of reproductive-age interviewed during 2015/16 KIHBS. When graphing method was used there was no strong preference for digit 0 and 5 as compared to the other digits as the determined peaks and troughs were also present in those other digits and this analogy was supported by Whipples index which gave the value of the index as 49.0 and 57.5 for terminal digit 0 and 5, respectively. This implied that the data was highly accurate and did not require adjustment before use. In addition, Myer's blended method was reported to be 2.9, which indicates there was insignificant preference and avoidance of ages for different terminal digits.

When data evaluation was undertaken on children ever born, it showed that majority of women reported to have ever given birth to 3 or less children, and the distribution of children ever born decreases when the number of children increases. The age distribution of these women with zero parities did not show any bias and was represented across all the ages in the reproductive group. The proportion of women with zero parity was insignificant and were therefore not included in the analysis.

KIHBS data showed some over reporting of parities, for example, some women reported children ever born that was biologically impracticable as there were some women aged 17 who reported to have had 12 children. Such unrealistic cases that were reported were not used in the estimation of fertility since they would distort the estimates.

Both reverse survival and OCMs method can produce current and reconstructed TFR estimates for upto 15 years before a survey. In the fourth and fifth year of the survey, reverse survival tends to overestimate fertility but the two techniques are consistent when the estimates are closer to the survey period.

### 5.3 Conclusion

Based on the results, this study concluded that both reverse survival and own children methods can use non birth history data and reliably estimate trends in fertility levels in Kenya. The two indirect methods can give consistent fertility estimates when the reference period is closer to the survey period but in the fourth and fifth year reverse survival method tends to systematically overstate fertility as compared to OCM and therefore the latter is preferred since it can give reliable and consistent current estimates and trends. Since estimates based on such methodologies compare with those from the full birth history method, it can be concluded that the non birth history data can be used to estimate current fertility and also produce a fertility trend for upto 15 years. When the two techniques were applied to the 2014 KDHS dataset they were found to under-report TFR compared to birth history method.

## 5.4 Recommendations

The results from this study put forth recommendations that would help Kenya obtain optimal fertility estimates from survey data.

### 5.4.1 Recommendations for policy

This study found out that in the absence of full birth history data, reverse survival and OCM methods can reliably estimate consistent fertility estimates and trend. Since FBH is known to require detailed data and hence is only applicable in fairly small samples and requires data that is collected periodically(Childs, 2019); (Avery et al., 2013) and (Collins and Levin, 2008), the study recommends that the Government should embrace the use of indirect techniques that utilize non birth history data to estimate fertility.

The 2015/16 KIHBS data did not collect data on recent births that are used in the estimation of fertility levels at national and county levels. This limitation could impact on the quality of lifetime fertility data as a result of omission of births due to recall lapse which is more pronounced in older women. The study recommends that future KIHBS should include questions on recent births in the questionnaire.

### 5.4.2 Recommendations for further study

This study encountered a few limitations. For example, reverse survival technique is recommended to be used only when generating national fertility estimates and does not estimate TFR by social characteristics, for example, education of mother. Since OCM technique can estimate

TFR at lower levels and subgroups then future studies should use it to get TFR estimates and trends by type of residence and by other social characteristics that affect fertility.

## Bibliography

Abbasi-Shavazi, M. (1997). An assessment of the own-children method of estimating fertility by birthplace in australia. *Journal of the Australian Population Association, Vol. 14, No. 2 (November 1997), pp. 167-185*.

Abbasi-Shavazi, P McDonald, M. H.-C. (2013). Assessment of the own–children estimates of fertility applied to the 2011 iran census and the 2010 iran-midhs. In *XXVII IUSSP International Population Conference, Busan, Korea, 26-31 August 2013*.

Abbasi-Shavazi, M. J. and McDonald, P. (2002). A comparison of fertility patterns of european immigrants in australia with those in the countries of origin. *Genus*, 58(1):53–76.

Avery, C., Clair, T., Levin, M., and Hill, K. (2013). The 'own children' fertility estimation procedure: A reappraisal. *Population Studies, 67(2), 171-183. https://doi.org/10.1080/00324728.2013.769616*.

Brass and Coale (1969). The demography of tropical africa. *Science*, 163(3863):168.

Caldwell, J. and Caldwell, P. (1987). The cultural context of high fertility in sub-saharan africa. *Population and Development Review, 13(3), 409-437. doi:10.2307/1973133*.

Casterline, J. B., Bongaarts, J., Askew, I., Maggwa, N., and Obare, F. (2017). Fertility Transitions in Ghana and Kenya: Trends, Determinants, and Implications for Policy and Programs. *Population and Development Review*, 43:289–307.

Childs, G. (2019). Demographic analysis of small populations using the own-children method. *Field Methods*, 16(4):379–395.

Cho, L., Retherford, R., and Choe, M. (1986). The own-children method of fertility estimation. *www.popline.org, Honolulu, Hawaii, The East-West Center, 1986. xvi, 188 p.*

Cho, L. J., W. H. G. and Bogue, D. J. (1970). Differential current fertility in the united states. *Chicago: University of Chicago*.

Cleland, J. (1996). Demographic data collection in less developed countries 1946-1996. *Population Studies*, 50(3):433–450.

Coale, A. and Trussell, J. (1996). The development and use of demographic models. *Population Studies*, 50(3):469–484.

Collins, O. and Levin, M. J. (2008). Fertility levels, trends and differentials in kenya: How does the own-children method add to our knowledge of the transition? *African Population Studies*.

Dubuc, S. (2009). Application of the own-children method for estimating fertility by ethnic and religious groups in the uk. *Journal of Population Research*, 26(3):207–225.

Dugbaza, T. . A. B. o. S. D. S. (1994). Recent trends and differentials in aboriginal and torres strait islander fertility, 1986-1991. *Canberra : Australian Bureau of Statistics, Demography Section, 1994*.

Garenne, M. and McCaa, R. (2017). 4-parameters own-children method: A spreadsheet for calculating fertility rates from census microdata: Application to selected african countries, in: Children, mothers and measuring fertility. *New Perspectives on the Own Child Method. Presented at the Cambridge Meeting, 18-20 September 2017, Cambridge, pp. 1–38*.

Hlabana, T. (2010). Application of the p / f ratio method in estimating fertility levels in lesotho.

Jain, S. (1989). Estimation of aboriginal fertility, 1971-86: An application of the own children method of fertility estimation. *Canberra : Australian Bureau of Statistics*.

KDHS (2014). Kenya Demographic and Health Survey 2014. Technical report, Rockville, MD, USA.

KenyaVision.2030 (2007). *Kenya vision 2030*. Government of the Republic of Kenya, Ministry of Planning and National Development and the National Economic and Social Council (NESC), Office of the President, Nairobi.

Kinfu, Y. (1993). Fertility trend in ethiopia: An application of 'own-children' method.

Kpedekpo, G. M. K. (1982). *Essentials of demographic analysis for Africa*.

Little, R. and Rubin, D. (1987). Statistical analysis with missing data. *John Wiley & Sons, New York.*

Mislevy, R. J., Rubin, D. B., and Little, R. J. A. (1991). Statistical analysis with missing data. *Journal of Educational Statistics*, 16(2):150–155.

Moultrie, Dorrington, H. H. T. . Z. T. f. D. E. T. M. R. D. A. H. K. H. I. T. and Zaba, B. (2013). Tools for demographic estimation. *2013 International Union for the Scientific Study of Population (IUSSP) 3-5 Rue Nicolas, Paris Cedex 20, France*.

Muhwava, W. and Timæus, I. M. (1996). Fertility decline in zimbabwe.

NCPD (2000). Sessional paper no. 1 of 2000 on national population policy for sustainable development. Technical report, National Council for Population and Development.

Ndagurwa, P. and Odimegwu, C. (2019). Small area estimation of fertility: Comparing the 4-parameters own-children method and the poisson regression-based person-period approach. *Spatial Demography*.

Preston, S. H., H. P. . G. M. (2001). Demography: Measuring and modeling population processes. *Oxford: Blackwell Publishers Ltd.*

Shryock and Siegel (1976). The methods and materials of demography. *Academic Press, San Diego.*

Spoorenberg, T. (2014). Reverse survival method of fertility estimation: An evaluation. *Demographic Research*, 31:217–246.

Timæus, I. and Moultrie, T. (2012). Estimation of fertility by reverse survival. *Tools for Demographic Estimation. Paris: International Union for the Scientific Study of Population*.

United Nations (1983). Manual ten. *Population studies ; no. 81*.

United Nations (2014). *Principles and Recommendations for a Vital Statistics System*. https://www.un-ilibrary.org/content/publication/63cf3eba-en.

Zlotnik, H. and Hill, K. (1981). The use of hypothetical cohorts in estimating demographic parameters under conditions of changing fertility and mortality. *Demography*, 18(1):103–122.