ESTIMATION OF FERTILITY IN KENYA: AN APPLICATION OF THE RELATIONAL GOMPERTZ AND REVERSE SURVIVAL MODELS

BY
ONG’ARO MOSORORI MOSES
Q56/80269/2012

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 2014
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.

Signature: ______________                      Date: ______________
Name: Ong’aro Mosorori Moses
Registration: Q56/80269/2012

This research project has been submitted for examination with our approval as University Supervisor(s):

Signature: ______________                      Date: ______________
Prof. Alfred O. Agwanda

Signature: ______________                      Date: ______________
Dr. Andrew Mutuku
DEDICATION

This research project is dedicated to Nelson Matieso, my late grandfather, for his kindness and wisdom; his selflessness will always be remembered.
ACKNOWLEDGEMENTS

It would not have been possible to write this Masters project without the help and support of the kind people around me, to only some of whom it is possible to give particular mention here. Above all, I would like to thank my wife Sophy for her personal support and great patience at all times. My parents Charles and Joyce Ong’aro for their financial support, brothers and sisters have given me their explicit support throughout, as always, for which my mere expression of thanks does not suffice.

This project would not have been possible without the help, support and patience of my principal supervisor, Prof. Otieno Alfred T.A, not to mention his advice and brilliant comments. The good advice, support and friendship of my second supervisor, Dr. Mutuku Andrew Kyalo, has been invaluable on both academic and personal level, for which I am extremely grateful. I would also like to thank all staff and teachers in the Population studies and research institute (PSRI) especially the administrator, Mr. Edward Sibota for his help in registration and providing guidance. Thanks are to God for giving me grace and mercies throughout this study.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEB</td>
<td>Children ever born</td>
</tr>
<tr>
<td>MCEB</td>
<td>Mean children ever born</td>
</tr>
<tr>
<td>DHS</td>
<td>Demographic Health Survey</td>
</tr>
<tr>
<td>GOK</td>
<td>Government of Kenya</td>
</tr>
<tr>
<td>ROK</td>
<td>Republic of Kenya</td>
</tr>
<tr>
<td>CBS</td>
<td>Central Bureau of Statistics</td>
</tr>
<tr>
<td>GFR</td>
<td>General Fertility Rate</td>
</tr>
<tr>
<td>KDHS</td>
<td>Kenya Demographic Health Survey</td>
</tr>
<tr>
<td>KNBS</td>
<td>Kenya National Bureau of Statistics</td>
</tr>
<tr>
<td>GSS</td>
<td>Ghana Statistical Services</td>
</tr>
<tr>
<td>PSRI</td>
<td>Population Studies Research Institute</td>
</tr>
<tr>
<td>TFR</td>
<td>Total Fertility Rate</td>
</tr>
<tr>
<td>ASFR</td>
<td>Age Specific Fertility Rate</td>
</tr>
<tr>
<td>KPHC</td>
<td>Kenya Population and Housing Census</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS

DECLARATION ......................................................................................................................... ii
DEDICATION ........................................................................................................................... iii
ACKNOWLEDGEMENTS ........................................................................................................ iv
ABBREVIATIONS .................................................................................................................. v
TABLE OF CONTENTS ........................................................................................................... vi
LIST OF TABLES .................................................................................................................... ix

## CHAPTER ONE: INTRODUCTION .................................................................................. 1
1.1 Background to the Study ................................................................................................. 1
1.2 Problem Statement ......................................................................................................... 3
1.3 Objectives of the Study ................................................................................................. 4
1.4 Justification of the study ............................................................................................... 4
1.5 Scope and Limitations of the Study .............................................................................. 5

## CHAPTER TWO: LITERATURE REVIEW ................................................................ 7
2.1 Introduction .................................................................................................................... 7
2.2 Applications of Relational Gompertz .......................................................................... 7
2.3 Applications of Reverse Survival ................................................................................ 9
2.4 Summary of literature review ...................................................................................... 11
2.5 Analytical framework ................................................................................................ 11
   2.5.1 The relational Gompertz model .............................................................................. 12
      2.5.1.1 Data required .................................................................................................. 15
      2.5.1.2 Assumptions ................................................................................................. 15
   2.5.2 Reverse survival .................................................................................................... 15
      2.5.2.1 Data required ................................................................................................ 16
      2.5.2.2 Assumptions ................................................................................................. 17
   2.5.3 Brass consistency check ....................................................................................... 17

## CHAPTER THREE: METHODOLOGY ...................................................................... 18
3.1 Introduction .................................................................................................................... 18
3.2 Source of data .............................................................................................................. 18
3.3 Methods of data analysis ............................................................................................ 18
   3.3.1 Relational Gompertz model ............................................................................... 19
LIST OF FIGURES

Figure 4.1 Age distribution in single years (2009 census) ................................................................. 33
Figure 4.2 Myers index for digit preference. ....................................................................................... 34
Figure 4.3 Distribution of reported age by 5 year age groups. ............................................................. 35
Figure 4.4 Distribution of adjusted female population by 5 year age groups .......................................... 36
Figure 4.5 Proportion not stated against proportion childless in each age group. ............................... 39
Figure 4.6 Total fertility trend by reverse survival method (1997-2009) for both adjusted and un adjusted TFRs .................................................................................................................. 42
Figure 4.7 Differentials of fertility by region (Relational gompertz model) ........................................... 43
Figure 4.8 Fertility by level of education ............................................................................................ 46
LIST OF TABLES

Table 4.1 Reported parities and proportion for unknown parity (u) .......................................................... 37
Table 4.3 Distribution of average parities, age specific fertility rates and Total fertility nationally .......... 41
Table 4.4 Fertility estimates per year(1995-2009).................................................................................. 41
Table 4.5 Fertility in rural Kenya by relational gompertz model............................................................ 44
Table 4.6 Fertility in urban Kenya by relational gompertz model........................................................... 44
Table 4.7 Fertility level of Never attended ........................................................................................... 45
Table 4.8 Fertility level of Primary......................................................................................................... 45
Table 4.9 Fertility level of Secondary and above .................................................................................. 45
Table 4.10 TFR estimates by various models......................................................................................... 46
ABSTRACT

Fertility in Kenya has been declining since 1980s. However, this decline experienced a stall during the period 2000-2007. The objective of this study was to determine levels and trends of fertility in Kenya. Specifically, this study sought to determine; levels and trends of fertility in Kenya, levels of fertility by region and fertility level by place of residence and education level. The study utilized 2009 census data.

To achieve the above objectives, the study utilized the relational gompertz and reverse survival models in estimating Kenyan fertility. Data quality was assessed by use of graphs, Whipples and Myers indices which indicated that the data was of poor quality and needed adjustment before use. This was done using 5 point method, 3 point moving average and El badry correction.

Estimates from the Relational gompertz model revealed that fertility in Kenya was 4.8 children per woman if the observed Age specific fertility rates applied until the end of child bearing period (15-49 years). The results also revealed that fertility had reduced in every province except Northeastern where fertility increased from 6.5 to 8.0 children per woman. Central province had experienced the utmost decline from 5.5 to 4.0 in 1999 and 2009 censuses respectively. The findings also revealed that women with secondary and above level of education had higher fertility compared to those with none with a TFR of 7.0, 5.8 and 3.9 for never attended school, primary and secondary and above respectively. Women living in rural areas had higher fertility compared to their urban counterparts with a TFR of 6.0 and 4.1 respectively. Estimates of TFR from the Reverse survival method revealed that fertility in Kenya was 4.3 children per woman with a stall in fertility decline during the period 200-2007 with a TFR of 4.9 children per woman in the reproductive age (15-49).

Based on the findings of this study, women with no education were found to have the highest TFR (7.0), there is therefore, need to increase the proportion of women with secondary and above education in Kenya as this will greatly reduce the level of fertility in Kenya. Although reverse survival estimates were consistent, the method tends to underestimate fertility levels, to improve the results; this study recommends that future studies should consider constructing life tables to be used as the standard because they will reflect better levels of mortality compared to the standard life tables used in this study.
1.1 Background to the Study

In Kenya, fertility increased rapidly during the 1960s and 1970s, culminating in the highest ever recorded total fertility rate (TFR) of 8.0 in 1979. The increase was attributed to improvement in the standard of living, low contraceptive use, low age at marriage and high value accorded to children. These demographic trends posed diverse challenges to the Government in so far as provision of basic needs was concerned (CBS, 2002b). TFR declined during the 1980s and 1990s changing from a high of 8.1 children per woman in the late 1970s to 6.7 in the late 1980s and dropping to 4.7 during the last half of the 1990s. However, fertility seemed to rise, albeit marginally, after 1998, reaching a TFR of 4.9 children per woman during the 2000-02 period. The TFR then declined, reaching a low of 4.6 children per woman during the 2006-08 period (KNBS and ICF macro 2010).

The Government's concern over the rapidly rising population growth rate in the 1960s and 1970s stimulated the adoption of policy strategies that laid the foundation for the onset of fertility transition in the late 1980s. For example, the Government officially adopted family planning policy in 1967 by establishing a maternal child health and family planning programme (MCH/FP) in the Ministry of Health. To further consolidate policy strategies, the Government established in 1982 the National Council for Population and Development (NCPD), a unit in the then Ministry of Planning and National Development (CBS, 2002b). In 2012, the government of Kenya passed a landmark policy to manage its rapid population growth. The new population policy aims to reduce the number of children a woman has over her lifetime from 5 in 2009 to 3 by 2030. The policy also includes targets for child mortality, maternal mortality, life expectancy, and other reproductive health measures (NCPD, 2012).

Kenya, like many other countries in sub-Saharan Africa, does not have a complete, reliable and accurate vital registration system. It is estimated that about 40% of births are registered (RoK, 1996:3). This problem has necessitated the use of demographic surveys and censuses to collect data on lifetime and current fertility. The indices of fertility that are normally calculated from censuses and surveys include Total fertility rate (TFR), Age specific fertility rate (ASFR) and
Crude birthrate (CBR). Such direct measures give unreliable values because of errors of omission and commission. Making comparisons between lifetime and current fertility sometimes enables such data to be adjusted (CBS, 2002b).

In order to reduce substantial errors inherent in such direct estimates of fertility levels and trends, it is recommended that indirect methods based on various techniques of data graduation be used (United Nations, 1983). The suitability of indirect methods, however, depends on the assumptions made about the nature of the data collected and the procedure used in their computation. For most developing countries such as Kenya, the assumptions often made have proved to be unrealistic and inconsistent with the changing demographic conditions. This has thus rendered the results of some indirect estimates of fertility levels liable to bias (CBS, 2002b). The Gompertz model has several advantages: for example, besides smoothing irregularities and providing the required ASFR, it also minimizes the effects of possible errors in the parities reported by older women (RoK, 1996:17).

It will also be important to note that, notwithstanding its simplicity, the reverse survival method of fertility estimation has seldom been applied. However, the method can be applied to a large body of existing and easily available population data, which has remained largely under-exploited. In contexts where limited demographic data are available the sole reliance of the method on age and sex distribution makes it of prime interest (Spoorenberg, 2014). Furthermore, the comparison of reverse survival fertility estimates with alternative fertility figures from multiple data sources (vital registration, sample surveys) provides an additional tool to evaluate the quality of population data. At a time of increasing reliance on sample survey data to estimate fertility for less statistically developed countries, the reverse survival method allows revisiting and highlighting the contribution of both historical and contemporary population census data to the study of fertility (Spoorenberg, 2014). Given the various highlighted strengths of the two methods, this study employed the use of reverse survival and the relational gompertz models of fertility estimation to estimate fertility in Kenya based on the 2009 census data.
1.2 Problem Statement

In order to study changes in fertility levels and trends in less developed countries, demographers have developed a series of estimation techniques based on data from census counts and household surveys (Brass 1975, Moultrie et al. 2012, United Nations 1983). Since the launch of the World Fertility Surveys (WFS) program in the late 1970s (and especially since the implementation of the Demographic and Health Surveys (DHS) program), population specialists have mostly relied on the ever-growing body of household surveys that have collected full birth history to derive fertility estimates in these countries. Yet, as a recent study has shown (Avery et al, 2013), alternative methods of fertility estimation can return very consistent fertility estimates using only basic demographic information.

Among the existing methods of fertility estimation, the reverse survival method is one of the most economical. Based on population data by age and sex collected in one census or single-round survey, the method consists in ‘reverse surviving’ those no longer present in the population of a given age in order to derive the number of births that occurred n years ago, using a set of probabilities of child and adult survivorship and age-specific fertility rates (Spoorenberg, 2014). The reverse survival method of fertility estimation is very similar to the own-children method of fertility estimation (Cho et al, 1986), but its data requirement is even lower.

Alternatively, the relational gompertz model, which is a refinement of the Brass P/F ratio that seeks to estimate age-specific fertility by determining the shape of fertility schedule from data on recent births reported in census or surveys while determining its level from the reported average parities of younger women. In producing estimates of age-specific and total fertility, the method seeks to remedy the errors commonly found in fertility data associated with too few or too many births being reported in the reference period, and the under-reporting of lifetime fertility and errors of age reporting among older women (Timæus and Moultrie, 2012).

Kenya, like many other countries in sub-Saharan Africa, does not have a complete, reliable and accurate vital registration system. It is estimated that about 40% of births are registered (RoK, 1996:3). This problem has necessitated the use of demographic surveys and censuses to collect data on lifetime and current fertility. The indices of fertility that are normally calculated from censuses and surveys include Total fertility rate (TFR), Age specific fertility rate (ASFR) and
Crude birthrate (CBR). Such direct measures give unreliable values because of errors of omission and commission. Making comparisons between lifetime and current fertility sometimes enables such data to be adjusted (CBS, 2002b). The estimates of fertility from this study were used to check the accuracy of estimates from surveys. In addition, estimates from reverse survival are consistent and accurate but the method is rarely used in Kenya. This study therefore, seeks to find out whether the estimates from reverse survival and relational gompertz models are consistent. This inquiry therefore, sought to answer the following research questions:

i) What are the levels of fertility in Kenya and by former provinces?

ii) How do the estimates from the two methods compare with that from the Brass consistency method?

1.3 Objectives of the Study

The general objective of this study was to estimate the levels of fertility in Kenya and by former provinces. The specific objectives were:

i) To determine fertility levels and trends in Kenya.

ii) To determine the level of fertility by region.

iii) To determine the level of fertility in Kenya by place of residence (rural and urban) and by education level.

1.4 Justification of the study

Fertility study is vital in demographic analysis because births are a central component of population growth. Studies in fertility help in checking on the progress of available population programs and interventions. In other words, they are used to assess the impact of interventions. TFR is one of the most important fertility indices that can be used to assess the impact of population programs, therefore, its measurement is important.

Among the existing methods of fertility estimation, the relational gompertz model, which is a refinement of the Brass P/F ratio that seeks to estimate age-specific fertility by determining the shape of fertility schedule from data on recent births reported in census or surveys while determining its level from the reported average parities of younger women. In producing estimates of age-specific and total fertility, the method seeks to remedy the errors commonly found in fertility data associated with too few or too many births being reported in the reference
period, and the under-reporting of lifetime fertility and errors of age reporting among older women (Timæus and Moultrie, 2012).

Alternatively, the reverse survival method is one of the most economical. Based on population data by age and sex collected in one census or single-round survey, the method consists in ‘reverse surviving’ those no longer present in the population of a given age in order to derive the number of births that occurred n years ago, using a set of probabilities of child and adult survivorship and age-specific fertility rates (Spoorenberg, 2014). The reverse survival method of fertility estimation is very similar to the own-children method of fertility estimation (Cho et al, 1986), but its data requirement is even lower and it is rarely used.

With regards to policy, Kenya has a new population policy whose targets include among others; reduce the average number of children per woman from 5 in 2009 to 3 children by 2030; reduce the natural growth rate from 25 people for every 1000 population in 2009 to 15 people for every 1000 population by 2030, reduce the deaths among children below one year of age from 52 in 2009 to 25 deaths per 100 live births by 2030 and also increase the use of family planning methods by married women from 46% in 2009 to 70% by 2030. In order to assess our performance towards the realization of vision 2030, there is need to ascertain the levels and trends of fertility in Kenya. Findings from this study will add to the already available information on fertility levels and trends in Kenya. Since fertility varies by region and education level (CBS, 2002b), just to mention a few, findings were used to assess where program managers needs to intensify family planning services or maintain current level. Estimates of fertility from this study, were also used to check on the accuracy of estimates from surveys that were done close to the census period of 2009.

1.5 Scope and Limitations of the Study
This study used data from the 2009 census. The 2009 Kenya Population and Housing Census (2009 KPHC) was the seventh census to be conducted since 1948 and the 5th since independence. It was conducted from the night of 24th/25th to 31st August 2009. The planning and execution of the 2009 KPHC was spearheaded by the Kenya National Bureau of Statistics (KNBS) on behalf of the Government of Kenya – in accordance with the Statistics Act 2006. The theme for the census was “Counting our people for the implementation of Vision 2030”.


theme was deemed pertinent in order to respond to the greater demands for statistical information for monitoring the implementation of Kenya’s current development goals and other global initiatives such as the Millennium Development Goals (MDGs). Thus, planning of the 2009 KPHC took cognizance of the need to collect a variety of information for monitoring the achievement of these goals.

Census data is a valuable source of demographic data though in Africa, especially Sub-Saharan Africa, census data is usually faced with challenges. Firstly, age heaping is a situation where by there is digit preference of ages ending with 0 and 5 or those that end with even numbers. Evidence from the 1999 census reveals a high preference of digit 0 and 5 with a Myers index of 16.4 (CBS, 2002b). This study used data grouped in the five year age groups in order to smoothen the data.

Secondly, there is an inherent problem of women with parity zero (childlessness) and those women that don’t state their parities. When the 1999 census data was scrutinized, fertility data revealed that, enumerators apparently did not fill in zeros in appropriate columns whenever the respondents indicated that they had never given birth. This was therefore coded as not stated during data processing (CBS, 2002b). This study used the El badry correction to correct the average parities. Details on the application of this correction are discussed in chapter four under data quality of this project report.

Lastly, errors affecting the quality of lifetime fertility data are omission of births, which arise from recall lapse by older women, especially of dead children or children who left the household soon after birth. They can also be caused by wrongful inclusion of still births, late foetal deaths and foster children among children ever born. This problem was addressed by the use of the relational gompertz model and most importantly, these errors did not affect the reverse survival model since it is mainly affected by age and sex distribution of the population which is addressed in chapter four under quality of data. It is also important to mention that the reverse survival method usually performs well at national level (Timæus and Moultrie, 2012) hence in this study; estimation of fertility at sub-national and sub groups was done by the relational gompertz model.
CHAPTER TWO
LITERATURE REVIEW

2.1 Introduction
This chapter presents literature on various studies on estimation of fertility. It summarizes the empirical findings of a range of studies that have used the reverse survival and relational gompertz models of indirect estimation of fertility and discusses the analytical frame work for this study. First is the literature on relational gompertz, then Reverse survival.

2.2 Applications of Relational Gompertz
Estimates of age-specific fertility rates based on survey data are known to suffer down-bias associated with incomplete reporting. In addition, many undeveloped countries lack the accurate civil registration of births needed to monitor fertility levels and trends over time (Avery, Clair, Levin and Hill 2013). The El Badry correction for non-response (El Badry, 1961; United Nations, 1983, pp. 230-235) provides a possible method of handling this problem (Blacker, 2002) and the average parities used for the construction of the hypothetical cohorts were first adjusted by the El Badry method as revealed in Blacker’s paper entitled “Kenya’s fertility transition: how low will it go?” The estimates refined by fitting fertility models to the average parities of the hypothetical cohorts, thus smoothing irregularities and reducing the vulnerability of the estimates to errors in the reports for older women (Blacker, 2002). Blacker also used Brass’s relational Gompertz fertility model (Brass, 1981) for that purpose.

While estimating the fertility levels of the Igbo area of Eastern Nigeria using different Demographic estimation techniques with the goal of arriving at the most accurate estimate of fertility in the area, Odimegwu used some of the most promising demographic techniques available for deriving fertility measures from the census and survey data. The results showed that, although different techniques were used with different assumptions, the derived estimates were very close. The relational gompertz model tends to give lower values (Odimegwu, 1998). Among the methods he used were: Mean CEB (45-49), Coale - Trassell method, Coale-Deneny method *(((Pr2)/2)/Pr1), Brass-Rachad method (Pr2)((2Pr1/Pr4)4)*, Brass P/F method, Brass consistency check method, Relational Gompertz model(mean parities) and Brass Relational Gompertz
model (current fertility). However it is important to note that Odimegwu didn’t use the reverse survival method.

In a comparative study to investigate the trends and determinants of fertility in Kenya and Uganda, Blacker, et al, used data from previous censuses and demographic and health surveys. When estimating fertility levels from the 2002 Uganda census, they used Brass’s relational Gompertz fertility model to smoothen the irregularities in the data and eliminate distortions liable to occur in the older age groups. The fitting procedure used was that developed by Zaba and expounded by Brass 1981. The model provides them with relative average parities, that is, the proportions of the total fertility achieved by women in the different age groups. Division of observed parities by the relative values yielded a series of implied fertilities, the consistency of which provided a measure of validity of the estimates. Indeed the consistencies of the estimates were remarkable, not only increasing the confidence of the quality of the data, but also strengthening their belief that the level of fertility in Uganda had remained virtually stable for some time (Blacker, et al, 2005).

The census data on lifetime and current births can be used in a variety of ways to examine fertility trends. Probably the simplest, and in some ways the most satisfactory, procedure, is the construction of “hypothetical cohorts” for the inter-censal periods from the differences in the average parities for the same cohorts of women in consecutive censuses (United Nations, 1983, pp. 59-64). The method uses only the average parities recorded in the censuses, and makes no assumptions about the nature or patterns of the possible errors in the data. However the census data on children ever born is not subject to the same detailed probing as in the birth history surveys, and errors, both of omission and faulty inclusions, undoubtedly occur. There is generally a category of women who are “not stated” as to the numbers of children they had borne. Most of these women are probably childless, but some may have borne children who, for a variety of reasons, were not recorded.

Previously, Brass (1964 and 1968) proposed a series of adjustments of such data to reflect more appropriate levels of fertility through comparison with data on children-ever-born by age, a measure of cohort-specific cumulative fertility. His now widely-used Parity/Fertility or PF ratio
method makes a number of strong assumptions, which have been the focus of an extended discussion in the literature on indirect estimation.

The use of Relational Gompertz fertility models fitted to data on current fertility and average parities of a hypothetical cohort with fitted relational Gompertz models experienced a significant change on the results when applied on the 1999 Kenyan census data. It is emphasized that the differences in fertility yielded by the two models depends on the number of points fitted in the Gompertz model. Moreover, fitting the Gompertz model to the uncorrected average parities is also less valuable. Fitting Gompertz model to the uncorrected current births will also give lower values than when the Gompertz model is fitted to average parity corrected by the El Badry correction method. In this latter case, the fertility values tend to be higher. Before one decides on the best estimate of fertility, it is desirable to consider other evidence on fertility during the 1980s and 1990s (CBS, 2002b).

In view of the shifting of events, misreporting of ages and underreporting of events, which affect current fertility data, an attempt was made to use two indirect fertility estimation techniques (P/F ratio method and the Gompertz relational model) to estimate TFRs for comparison with the reported TFRs (GSS, 2013). The 917 births (0.15 percent) and 8,638 births (1.39 percent) that occurred to women aged 12-14 years and 50-54 years respectively were excluded from the calculation of the reported total fertility rates. The reported TFRs based on current fertility data appeared to be lower than the estimated TFRs, indicating possible under-reporting of births. The reported TFR for 2010 (3.28) is lower than the two estimated TFRs which are 4.71 and 4.57 derived with the P/F ratio method and Gompertz relational model respectively. Because the reported TFRs are low due in part to misstatement of age and underreporting of current fertility, Brass relational Gompertz model was used to compute TFRs for the regions and urban and rural areas for 2000 and 2010 Censuses in Ghana (GSS, 2013).

2.3 Applications of Reverse Survival
The reverse survival method produces total fertility estimates that are very consistent and hardly affected by erroneous assumptions on the age distribution of fertility or by the use of incorrect mortality levels, trends, and age patterns. The quality of the age and sex population data that is
‘reverse survived’ determines the consistency of the estimates. The contribution of the method for the estimation of past and present trends in total fertility is illustrated through its application to the population data of five countries characterized by distinct fertility levels and data quality issues namely: Japan, Algeria, Mongolia, Ghana, and Kenya (Spoorenberg, 2014). A simulated population was first projected over 15 years using a set of fertility and mortality age and sex patterns. The projected population was then reverse survived using the Excel template FE_reverse_4.xlsx, provided with Timæus and Moultrie (2012).

Reverse survival fertility estimates were then compared for consistency to the total fertility rates used to project the population. The sensitivity was assessed by introducing a series of distortions in the projection of the population and comparing the difference implied in the resulting fertility estimates. Notwithstanding its simplicity, the reverse survival method of fertility estimation has seldom been applied. The method can be applied to a large body of existing and easily available population data both contemporary and historical that so far has remained largely under-exploited, and contribute to the study of fertility levels and trends. For Kenya, the population by single age and sex from the last five censuses (i.e., 1969, 1979, 1989, 1999, and 2009) was used to compute reverse survival fertility estimates (United Nations Statistics Division 2013b). While the patterns of under-enumeration of young children and age heaping were evident, the reverse survival fertility estimates were generally in line with alternative fertility figures (Spoorenberg, 2014).

The own-children method of fertility estimation tracks temporal changes in fertility patterns. Opiyo and Levin revisited the Kenyan fertility transition by applying the method to 1979, 1989 and 1999 censuses, and 1989, 1993, 1998 and 2003 Demographic and Health Surveys data. The method’s ability to provide yearly fertility rates for periods preceding each data source adds enormous knowledge to fertility patterns. For Kenya, these trends went back through the 1960s. First, the method sheds additional light on the onset of the transition. Second, the trends highlight major differences in the onset and pace of fertility decline among regions and key sub-groups. Third, the rates for overlapping periods provide both internal and external validity checks that heighten confidence in the overall results. Last, it provides a rare opportunity to evaluate birth
history fertility rates. They then concluded that “taken together, these estimates provided more
detail than ever before regarding fertility patterns in Kenya” (Opiyo and Levin, 2008).

Three methods of estimating birth rate under non stable conditions were examined. The first is
the Coal’s robust estimation of birth rate using proportion under age 15 (both sexes) and the
mortality estimate l5 obtained by the Brass type indirect estimation. The second is the method
based on the generalized stable population equation developed by Preston and Coale using the
data on age-sex distributions for two recent censuses and l5. The third method is the well known
reverse survival ratio method using the proportion under age 15 and l5 for the recent census.
These three methods are shown to belong to a broad system of reverse survival methods. It was
also shown that they all give robust estimates and in most cases give extremely close values for
birth rates (Venkatacharya and Teklu 1987).

2.4 Summary of literature review
Data collected from the census however rich, is faced with challenges. First, there is inherent
omission of births by older women due to recall lapse. Second, there is a problem of women
whose parity is not stated. There is an apparent trend of enumerators recording 0 for parity even
when respondents have not stated their parity or remained silent. This problem can be corrected
by the application of the El badry correction procedure before using the relational Gompertz
model to smoothen the irregularities in data and distortions liable in the older age groups (CBS
2002; Blacker, et al, 2005 and GSS, 2013). The estimates obtained when the relational gompertz
model is fitted to data on average parities after an El badry correction have been consistent and
accurate. On the other hand, estimates from the reverse survival model are robust and very close
to values obtained by other methods. Despite its economic nature of requiring data on age and
sex and ability to produce levels and trends of fertility, it is rarely used, instead it’s extension
‘the own children’ has been.

2.5 Analytical framework
Data on fertility from the 2009 Kenyan census was analyzed using two fertility models: The
relational Gompertz model and the reverse survival model. The estimates from the two models
were compared with those from the Brass consistency check for consistency. This section
discusses the main assumptions of the models, data required and brief description. Much about
the models is covered in chapter three of this Project.

2.5.1 The relational Gompertz model
The relational Gompertz method is a refinement of the Brass P/F ratio method that seeks to
estimate age-specific and total fertility by determining the shape of the fertility schedule from
data on recent births reported in censuses or surveys while determining its level from the
reported average parities of younger women. In producing estimates of age-specific and total
fertility, the method seeks to remedy the errors commonly found in fertility data associated with
too few or too many births being reported in the reference period, and the under-reporting of
lifetime fertility and errors of age reporting among older women. The method relies on a useful
property of a (cumulated) Gompertz distribution, \(G(x) = \exp(a \cdot \exp(bx))\), which is sigmoidal
(i.e. S-shaped), but also has an associated hazard function that is right-skewed and which
therefore captures fairly well both the pattern of average parities of women by age and their
cumulated fertility. The form of \(G(x)\) implies that a double-negative log transform of
proportional cumulated fertilities or average parities approximates a straight line for most of the
age range. The double-log transform,

\[
Y(x) = -\ln(-\ln(G(x)))
\]

is termed a goompit and has a close analogue in the logit transform frequently used in mortality analysis. Brass, however, found that a much closer linear fit could be
obtained by a relational model that expresses the gompits of an observed series of fertility data as
a linear function of the gompits of a defined standard fertility schedule. In other
words, \(Y(x) = \alpha + \beta Y^*(x)\) Where \(Y^*(x)\) is the gompit of the standard fertility schedule. Evidently,
if \(\alpha = 0 \) and \(\beta = 1\), the fertility schedule will be identical to the standard fertility schedule. Alpha
(\(\alpha\)) represents the extent to which the age location of childbearing in the population differs from
that of the standard (negative values imply an older distribution of ages at childbearing than in
the standard), while beta (\(\beta\)) is a measure the spread of the fertility distribution (values greater
than 1 imply a narrower distribution).

As input data, the method requires average parities at each age group, \( sP_{15} \) for \( x = 15, 20, \ldots, 45 \),
and fertility rates in each age group, \( sf_{15} \). For ease of exposition, and to differentiate more clearly
between lifetime and recent fertility data, \( sP_{15} \) is indexed as \( P(1) \), \( sP_{20} \) as \( P(2) \) and so on.
The derivation of these inputs from census data is described in the section of the manual dealing with the assessment and evaluation of data on fertility. As with other methods, the average parities should be adjusted for an el-Badry correction where appropriate. Cumulated (period) fertility to the end-point of each age group, is given by, \[ F(x+5) = \alpha \sum_{i=5}^{x} g_i \]

The original method proposed by Brass (1978) used the series of the gompits of the ratio of cumulated fertility to the end of each age group to the fertility rate cumulated to age 50 (i.e. total fertility, TF), giving a sigmoidal curve with minimum of 0 and a maximum (at the last age group) of 1. Gompits of the average parities are derived in a similar manner.

There are two inherent weaknesses in this approach. First, it requires total fertility as an input, and estimates of total fertility available from reported age-specific fertility rates (ASFRs) may be biased. In fact, total fertility is often the parameter of greatest interest that the analyst is trying to estimate. The second weakness is the implicit assumption of constant fertility over time arising from the treatment of the parity gompits. Nonetheless, Brass’ formulation inspired the derivation of the standard fertility schedule by Booth (1980, 1984), which is still used in the model to this day.

Both limitations are addressed comprehensively by Zaba’s (1981) reformulation of the method, which avoids the circularity of the original method while also dropping the need to assume that fertility has been constant. Further unpublished work by Zaba generalized the approach to incorporate alternative variants of the model (some of which are described here). In summary, she showed that the model can be expressed as:

\[ z(x) - e(x) = \alpha + \beta g(x) + (c/2)(\beta-1)^2, \quad \text{Equation 1} \]

where \( e(x) \), \( g(x) \) and \( c \) are functions of the chosen standard and \( z(x) \) is the gompit of the ratios of adjacent cumulated period fertility measures, i.e. \( F(x)/F(x+5) \), instead of \( F(x)/50 \) as Brass originally suggested. In other words, \( Z(x) = -\ln(-\ln(F(x)/F(x+5))) \).

For the parity data, the model is fitted to the ratios of adjacent average parities, \( P(i)/P(i+1) \). This means that the model can be used without the need to estimate total fertility before fitting the shape parameters. It follows further from Equation 1 that a plot of \( z(x) - e(x) \) against \( g(x) \) should
be a straight line with slope $\beta$ and intercept $\alpha + (1/2)c(\beta-1)^2$. (Noting that $\beta$ should be close to one, early formulations of the procedure deemed the last term of the intercept unimportant, leaving the intercept approximated by alpha. With the computing power now to hand, there is no justification for the associated loss of precision in the calculation of the intercept. However, the requirement that $\beta$ be close to 1 remains).

Exactly the same reasoning holds for the evaluation of the parity data. Using $P(i)/P(i+1)$, the ratio of average parities in successive age groups, with a linear equation relating $z(i) - e(i)$ to $g(i)$ results in:

$$z(i) - e(i) = \alpha + \beta g(i) + c/2(\beta-1)^2.$$  

Equation 2

By convention, the points derived from the parity data are known as $P$-points and those derived from the fertility rates are known as $F$-points. The goal of the model-fitting procedure is to find a combination of $P$- and $F$-points that are internally consistent with each other (i.e. the two sets of points define essentially the same lines) and then to use these to determine jointly the parameters $\alpha$ and $\beta$ in Equations 1 and 2 above. The values of $\alpha$ and $\beta$ are used to derive the relational gompits, $Y(x) = \alpha + \beta Y'(x)$, and similarly for $Y(i)$.

Deriving a fitted fertility distribution using the relational Gompertz method requires tabulations of calculated average parities and fertility rates by age. The fertility rates are cumulated and ratios of successive cumulated values are computed. Ratios of successive average parities are also calculated. Gompits of these ratios are calculated and used to plot the two pairs of points, $z(x) - e(x)$ against $g(x)$, and $z(i) - e(i)$ against $g(i)$. The fitted lines will have slopes equal to $\beta$, and an intercept term involving $\alpha$, $\beta$ and $c$, from which $\alpha$ can be calculated. The values of $\alpha$ and $\beta$ are used to transform the gompits of the standard cumulants into fitted gompits, which are then converted to fitted average parities and fertility rates. The level of fertility is set by the most reliable parity points. These are usually those on women aged 20-29 or 20-34 who are both less likely to omit births and likely to report their ages more accurately than older women.

The use of the relational Gompertz model in the calculation of a fitted fertility distribution has a number of advantages over the earlier $P/F$ ratio method. The model uses a reliable fertility pattern for medium- to high-fertility regimes (the Booth standard). Thus unreliable fertility rates estimated from reports of births in the last year can be replaced by model values which are fitted using the more reliable points. The plot of the two series of points is a powerful guide to the
reliability of each point, and can provide insight into data errors as well as identify fertility trends. All reliable points can be used to derive the fitted model distribution. The model also provides a reliable way of interpolating between values to make parity and cumulated fertility data comparable and to convert fertility rates in unconventional age groups to rates that apply to conventional age groups.

2.5.1.1 Data required.
Fertility rates for the 12, 24 or 36 months before the survey, classified by age of mother at survey, or by age at birth of child; or number of women at the census or survey date, by five-year age group; and number of births to women in the 12, 24 or 36 months before the survey, by five-year age group. Average parities of women classified by five-year age group of mother; or number of women by five-year age group; and total number of children born to women, by five-year age group.

2.5.1.2 Assumptions
- The standard fertility schedule chosen for use in the fitting procedure appropriately reflects the shape of the fertility distribution in the population.
- Any changes in fertility have been smooth and gradual and have affected all age groups in a broadly similar way.
- Errors in the pre-adjustment fertility rates are proportionately the same among women in the central age groups (20-39), so that the age pattern of fertility described by reported recent births is reasonably accurate.
- The parities reported by younger women (aged 20-29 or 20-34) are accurate. The method usually allows violations of these assumptions to be detected.

2.5.2 Reverse survival
Reverse survival is a method for estimating fertility from data collected in a census or single-round survey that can be used even if no questions have been asked about fertility directly. In a population closed to migration, the population of any age \( x \) are the survivors of the births in that population \( x \) completed years previously. This implies that the number of births occurring \( x \) years ago can be calculated, provided that one can estimate the life table survival probabilities from
birth to age \( x \) (that is, \( L_x/\theta \) "Reverse surviving" the population to its birth year and dividing by an estimate of the total population in that year gives the crude birth rate, while dividing by an estimate of women of childbearing age gives the General Fertility Ratio. By combining reverse survival estimates of past births and women according to age with estimates of, or a reasonable assumption about, the age pattern of fertility, one can also estimate Total Fertility.

Asking about births in the last year or the date of women’s last live birth only provides an estimate of current fertility. In contrast, reverse survival methods can provide estimates of fertility for the last 15 years. Moreover, unlike fertility estimates from birth histories, which are usually collected only from women aged 15 to 49, fertility estimates produced by reverse survival do not become increasingly truncated at older ages as they are calculated for more distant periods. So long as a single-year age distribution of children is available, the approach can produce an annual series of fertility estimates. In practice though, the data on age collected in developing countries are seldom sufficiently accurate to yield an undistorted time series. The computational steps are covered in chapter 3 of this proposal.

### 2.5.2.1 Data required

To derive General Fertility Ratios for individual years, the following data are required:

- Tabulations of the population (of both sexes) aged 0 to 14, by single years of age.
- Tabulations of the female population aged 15 to 64 by five-year age group.
- Cohort survival probabilities, \( L_x \), for children aged 0 to 14 of both sexes.
- Survivorship ratios, \( 5L_x/5L_x \) for adult women for each of the three five-year periods preceding the inquiry.

In respect of the mortality estimates, the implementation of the method in the associated Excel workbook allows these to be specified either by reference to period-specific parameters \( \alpha \) and \( \beta \) of appropriate relational model life tables, or to identified values of \( 5q_0 \) (for children) and \( 45q_{15} \) (for adult women) for each of the three five-year periods preceding the enquiry.

To produce estimates of Total Fertility, one also requires either

- A single age-specific fertility distribution that is assumed to apply to the entire period covered by the estimates, or
Two age-specific fertility distributions, one of which applies to a date reasonably close to the index inquiry and the other to a date approximately 15 years prior to that.

Either a series of fertility rates or the parameters of a relational Gompertz model fitted to a standard fertility schedule can be used as an input to the calculations. Note that only the estimated shape of the fertility distribution is based on these fertility schedules. It is the estimated number of births relative to the population of women of childbearing age that almost entirely determines the estimates of the General and Total Fertility Ratios.

2.5.2.2 Assumptions
The population is assumed to have been closed to migration for as many years as are covered by the reverse survival estimates. However, because children usually migrate with their mothers, errors in the numerator and denominator of the estimated rates largely cancel out. Significant bias will result only if migration flows are large and migrants have different fertility from the rest of the population.

2.5.3 Brass consistency check
This method was developed by Brass to check for the accuracy of the TFR estimates as applied by. The Brass consistency check is given as:
\[
\frac{CT}{W} = F \left[ \frac{W - m}{n} + \frac{P}{2n} (m(n - m) - 40) \right]
\]

Where,

- \(W\) = Total number of women aged 15-44
- \(CT\) = Number of children ever born to women 15-44
- \(M\) = Mean age of fertility distribution minus 15, let \(f_i\) be the age specific fertility rate for the first age group i.e. 15-19, then \(m = [17 \times f_{15} + 22 \times f_{16} + 27 \times f_{17} + 32 \times f_{18} + 37 \times f_{19} + 42 \times f_{20}]\) and 17,22 ........42 are class midpoints.
- \(P\) = Average proportional change in the year in the number of women aged 15-44, given by,
\[
\frac{2(W_{30-39} - W_{15-30})}{15W}
\]
- \(F\) = Total fertility rate
- \(n\) = 30 years
CHAPTER THREE
METHODOLOGY

3.1 Introduction
This chapter describes the source of data and the methods that were utilized in data analysis.

3.2 Source of data
This study used data from the 2009 census. The 2009 Kenya Population and Housing Census (2009 KPHC) was the seventh census to be conducted since 1948 and the 5th since independence. It was conducted from the night of 24th/25th to 31st August 2009. The planning and execution of the 2009 KPHC was spearheaded by the Kenya National Bureau of Statistics (KNBS) on behalf of the Government of Kenya – in accordance with the Statistics Act 2006. The theme for the census was “Counting our people for the implementation of Vision 2030”. This theme was deemed pertinent in order to respond to the greater demands for statistical information for monitoring the implementation of Kenya’s current development goals and other global initiatives such as the Millennium Development Goals (MDGs). Thus, planning of the 2009 KPHC took cognizance of the need to collect a variety of information for monitoring the achievement of these goals. This study utilized 100% of the 2009 census data as required by the two methods. The KPHC collected information on: Composition (age and sex) of the population, fertility, mortality, nuptiality, migration, urbanization, housing conditions, availability of household amenities, distribution of persons with disability, labor force participation and levels of education attained by her population. This study required data on fertility and age as detailed in this chapter.

3.3 Methods of data analysis
The methods of data analysis were: The reverse survival model, the relational gompertz model and the Brass’ method of consistency check. Except the Brass consistency check, this study used the Excel templates provided by Timæus and Moultrie (2012) and the estimates were compared with those calculated from Brass consistency check. The description of the methods was adopted from Timæus and Moultrie (2012).
3.3.1 Relational Gompertz model

The relational Gompertz method is a refinement of the Brass P/F ratio method that seeks to estimate age-specific and total fertility by determining the shape of the fertility schedule from data on recent births reported in censuses or surveys while determining its level from the reported average parities of younger women. In producing estimates of age-specific and total fertility, the method seeks to remedy the errors commonly found in fertility data associated with too few or too many births being reported in the reference period, and the under-reporting of lifetime fertility and errors of age reporting among older women. The method relies on a useful property of a (cumulated) Gompertz distribution, $G(x) = \exp(a \exp(bx))$, which is sigmoidal (i.e. S-shaped), but also has an associated hazard function that is right-skewed and which therefore captures fairly well both the pattern of average parities of women by age and their cumulated fertility. The form of $G(x)$ implies that a double-negative log transform of proportional cumulated fertilities or average parities approximates a straight line for most of the age range. The double-log transform, $Y(x) = -\ln(-\ln(G(x)))$ is termed a gompit and has a close analogue in the logit transform frequently used in mortality analysis. Brass, however, found that a much closer linear fit could be obtained by a relational model that expresses the gompits of an observed series of fertility data as a linear function of the gompits of a defined standard fertility schedule. In other words, $Y(x) = \alpha + \beta Y^s(x)$ Where $Y^s(x)$ is the gompit of the standard fertility schedule. Evidently, if $\alpha = 0$ and $\beta = 1$, the fertility schedule will be identical to the standard fertility schedule. Alpha ($\alpha$) represents the extent to which the age location of childbearing in the population differs from that of the standard (negative values imply an older distribution of ages at childbearing than in the standard), while beta ($\beta$) is a measure the spread of the fertility distribution (values greater than 1 imply a narrower distribution).

As input data, the method requires average parities at each age group, $P_x$ for $x = 15, 20, \ldots, 45$, and fertility rates in each age group, $f_x$. For ease of exposition, and to differentiate more clearly between lifetime and recent fertility data, $P_{15}$ is indexed as $P(1)$, $P_{20}$ as $P(2)$ and so on. The derivation of these inputs from census data is described in the section of the manual dealing with the assessment and evaluation of data on fertility. As with other methods, the average parities should be adjusted for an el-Badry correction where appropriate. Cumulated (period) fertility to the end-point of each age group, is given by, $F(x+5) = \sum_{g=15}^{x} s_f$.
The original method proposed by Brass (1978) used the series of the gompits of the ratio of cumulated fertility to the end of each age group to the fertility rate cumulated to age 50 (i.e. total fertility, TF), giving a sigmoidal curve with minimum of 0 and a maximum (at the last age group) of 1. Gompits of the average parities are derived in a similar manner.

There are two inherent weaknesses in this approach. First, it requires total fertility as an input, and estimates of total fertility available from reported age-specific fertility rates (ASFRs) may be biased. In fact, total fertility is often the parameter of greatest interest that the analyst is trying to estimate. The second weakness is the implicit assumption of constant fertility over time arising from the treatment of the parity gompits. Nonetheless, Brass’ formulation inspired the derivation of the standard fertility schedule by Booth (1980, 1984), which is still used in the model to this day.

Both limitations are addressed comprehensively by Zaba’s (1981) reformulation of the method, which avoids the circularity of the original method while also dropping the need to assume that fertility has been constant. Further unpublished work by Zaba generalized the approach to incorporate alternative variants of the model (some of which are described here). In summary, she showed that the model can be expressed as:

\[ z(x) - e(x) = \alpha + \beta g(x) + \left( \frac{c}{2} \right) (\beta - 1)^2, \]  

Equation 1

where \( e(x), g(x) \) and \( c \) are functions of the chosen standard and \( z(x) \) is the gompit of the ratios of adjacent cumulated period fertility measures, i.e. \( F(x)/F(x+5) \), instead of \( F(x)/50 \) as Brass originally suggested. In other words, \( Z(x) = -\ln(-\ln(F(x)/F(x+5))) \).

For the parity data, the model is fitted to the ratios of adjacent average parities, \( P(i)/P(i+1) \). This means that the model can be used without the need to estimate total fertility before fitting the shape parameters. It follows further from Equation 1 that a plot of \( z(x) - e(x) \) against \( g(x) \) should be a straight line with slope \( \beta \) and intercept \( \alpha + (1/2)*c \cdot (\beta - 1)^2 \). (Noting that \( \beta \) should be close to one, early formulations of the procedure deemed the last term of the intercept unimportant, leaving the intercept approximated by alpha. With the computing power now to hand, there is no justification for the associated loss of precision in the calculation of the intercept. However, the requirement that \( \beta \) be close to 1 remains).
Exactly the same reasoning holds for the evaluation of the parity data. Using $P(i)/P(i+1)$, the ratio of average parities in successive age groups, with a linear equation relating $z(i) - e(i)$ to $g(i)$ results in:

$$z(i) - e(i) = \alpha + \beta g(i) + c/2(\beta - 1)^2, \quad \text{Equation 2}$$

By convention, the points derived from the parity data are known as $P$-points and those derived from the fertility rates are known as $F$-points. The goal of the model-fitting procedure is to find a combination of $P$- and $F$-points that are internally consistent with each other (i.e. the two sets of points define essentially the same lines) and then to use these to determine jointly the parameters $\alpha$ and $\beta$ in Equations 1 and 2 above. The values of $\alpha$ and $\beta$ are used to derive the relational gompits, $Y(x) = \alpha + \beta Y'(x)$, and similarly for $Y(i)$.

Deriving a fitted fertility distribution using the relational Gompertz method requires tabulations of calculated average parities and fertility rates by age. The fertility rates are cumulated and ratios of successive cumulated values are computed. Ratios of successive average parities are also calculated. Gompits of these ratios are calculated and used to plot the two pairs of points, $z(x) - e(x)$ against $g(x)$, and $z(i) - e(i)$ against $g(i)$. The fitted lines will have slopes equal to $\beta$, and an intercept term involving $\alpha$, $\beta$ and $c$, from which $\alpha$ can be calculated. The values of $\alpha$ and $\beta$ are used to transform the gompits of the standard cumulants into fitted gompits, which are then converted to fitted average parities and fertility rates. The level of fertility is set by the most reliable parity points. These are usually those on women aged 20-29 or 20-34 who are both less likely to omit births and likely to report their ages more accurately than older women.

The use of the relational Gompertz model in the calculation of a fitted fertility distribution has a number of advantages over the earlier $P/F$ ratio method. The model uses a reliable fertility pattern for medium- to high-fertility regimes (the Booth standard). Thus unreliable fertility rates estimated from reports of births in the last year can be replaced by model values which are fitted using the more reliable points. The plot of the two series of points is a powerful guide to the reliability of each point, and can provide insight into data errors as well as identify fertility trends. All reliable points can be used to derive the fitted model distribution. The model also provides a reliable way of interpolating between values to make parity and cumulated fertility data comparable and to convert fertility rates in unconventional age groups to rates that apply to conventional age groups.
3.3.1.1 Data required.
Fertility rates for the 12, 24 or 36 months before the survey, classified by age of mother at survey, or by age at birth of child; or number of women at the census or survey date, by five-year age group; and number of births to women in the 12, 24 or 36 months before the survey, by five-year age group. Average parities of women classified by five-year age group of mother; or number of women, by five-year age group; and total number of children born to women, by five-year age group.

3.3.1.2 Assumptions
- The standard fertility schedule chosen for use in the fitting procedure appropriately reflects the shape of the fertility distribution in the population.
- Any changes in fertility have been smooth and gradual and have affected all age groups in a broadly similar way.
- Errors in the pre-adjustment fertility rates are proportionately the same among women in the central age groups (20-39), so that the age pattern of fertility described by reported recent births is reasonably accurate.
- The parities reported by younger women (aged 20-29 or 20-34) are accurate. The method usually allows violations of these assumptions to be detected.

3.3.1.3 Application of the method
The method is applied in the following stages.

**Step 1: Calculate the reported average parities**
Calculate the average parities, \( \bar{P}_x \) of women in each age group \( [x, x+5) \), for \( x = 15, 20 \ldots 45 \), if not already done as part of the preliminary investigations, or produced as a consequence of applying the el-Badry correction.

**Step 2: Determine the classification of the age of mother**
Depending on the data available, the fertility rates may be classified either by age of mother at the survey date, or by age of mother at birth of her child. The former ages are almost always encountered in the analysis of census data, where the mother’s age is her age at the census. The latter are more commonly encountered with administrative data derived from vital registration
systems. It is crucial that this classification is determined correctly as mis-specification here will bias the estimated rates produced.

The spreadsheet implementation of the model can accommodate data with no shift (i.e. reported according to the age of mother at birth); or – in the case of data classified by age of mother at survey date – with half a year, a year or one and a half year’s shift (for periods of investigation of 12, 24 and 36 months, respectively).

**Step 3: Calculate implied age-specific fertility rates and parities**

Age-specific fertility rates are derived by dividing the births reported in the period of investigation (e.g. the year, two years or three years) before the survey date by the number of women in each age group.

**Step 4: Choose the fertility standard to be applied and the model variant to be fitted**

The default fertility standard is that produced by Booth, modified slightly by Zaba (1981). The standard is appropriate to high- and medium-fertility populations and is a normalized cumulated fertility schedule (i.e. with total fertility equal to one). The standard $Y_s(x)$ values are determined by taking the gompits of the schedule and the standard parity values, $Y_s(i)$, are the gompits of the parities associated with the standard fertility schedule. The choice of standard determines the values of $g()$ and $e()$ used in the regression fitting procedures which are derived algebraically from the $Y_s()$.

Two variants of the relational Gompertz model are presented here. The default option is to make the same assumptions about the nature of errors inherent in fertility data as in the Brass $P/F$ method, namely that reports of recent fertility suffer from reference-period errors and under-reporting that are independent of age, and that reports of lifetime fertility suffer from omission errors that increase with age. In the spreadsheet, this is referred to as the 'Shape $F$ – Level $P$' variant.

A second variant involves using the relational Gompertz model to correct for possible distortions in the *shape* of the fertility distribution, while leaving the level unchanged. Clearly, if reference period errors or under-reporting are suspected, this variant will not give a plausible estimate of fertility.
Step 5: Evaluate the plot of P-points and F-points

The plots of $z(x) - e(x)$ against $g(x)$, and $z(i) - e(i)$ against $g(i)$ on the same set of axes are then used as a diagnostic for identifying common errors and trends in the data.

Step 6: Fit the model by selecting the points to be used

Initially, all points should be included in the model, the only exception being if the average parities in one age group are higher than the average parities in the next. In this case the gompit will be undefined and the model cannot be fitted using that point. (Such a situation cannot occur in a real cohort, but could arise because of data error or in a synthetic cohort during a time of rapidly changing fertility.)

If the parity and fertility data are internally consistent, the plots of $z() - e()$ against $g()$ should result in straight lines. Those $P$-points and $F$-points that cause each plot to deviate from a straight line should be excluded from the model. Ordinary least squares regression is used to fit lines to the $P$-points and $F$-points and to identify, sequentially, those points that do not fit neatly on a straight line. The intention is to seek the largest combination of $P$- and $F$-points that lie (almost) on the same line, and to use these to fit the model.

Points are selected for inclusion or exclusion using the following guidelines: A contiguous series of points must be included in the model. Sequentially, only the end-most points can be excluded. (The reason for this is that each point on the graph is the result of calculations involving the ratio of a pair of adjacent data values. If the analysis leads you to conclude that a data value is unreliable as a denominator of one of these ratios, it is not logical to accept it as the numerator of the next ratio.)

$P$-points should be eliminated in preference to $F$-points. This is because the average parity data are generally more prone to age-specific errors than the fertility data. $P$-points which deviate clearly from the straight line based only on the other $P$-points as well as $F$-points which deviate clearly from the straight line based only on the other $F$-points should be eliminated early on in the fitting process. $P$- and $F$-points at older ages should be eliminated in preference to those at younger ages since data at these ages are usually the least reliable and show the least consistency between lifetime and recent fertility. The exception to this relates to the data points for women
under the age of 20 because small numbers of events, as expected for younger women, frequently make the estimates of average parities or cumulated fertility unreliable.

Where only a marginally worse fit is achieved with more points, this is to be preferred to a slightly better fit achieved with fewer points. The spreadsheet calculates the root mean squared error (RMSE)

\[ RMSE = \sqrt{\frac{\sum (z(i) - s(i))^2}{n}} \]

from the points used to fit the model. This statistic can assist with determining the optimal number of data points to which to fit if there is uncertainty as to which of two competing models is better. In this situation, one can choose the model with the lower RMSE.

**Step 7: Assess the fitted parameters**

The values of \( \alpha \) and \( \beta \) that represent the best-fitting line joining the remaining \( P \)-points and \( F \)-points must be checked to confirm that they are not so far from their central values as to suggest that the standard chosen is inappropriate. A good fit is indicated if \(-0.3 < \alpha < 0.3\), and if \( 0.8 < \beta < 1.25 \). If the parameters lie outside this range, one or both of the underlying data series are problematic or the standard is inappropriate. Experimentation with another standard (see below) or changing the selection of points should be done before proceeding further. If the parameters still lie outside the ranges above, the method should be regarded as inappropriate.

**Step 8: Fitted ASFRs and total fertility**

Having estimated the two parameters of the model, they can be applied to the standard values for the parities to obtain fitted values \( Y(i) = \alpha + \beta Ys(i) \). These are then converted back into measures of the cumulative proportion of fertility achieved by age group \( i \) using the anti-gompit transformation. The anti-gompits based on the parity distributions indicate the proportion of fertility achieved by that age group. Dividing the observed parity in each age group by these proportions produces a series of estimates of total fertility. Averaging these values across the sub-set of age groups that were used to estimate \( \alpha \) and \( \beta \) gives the fitted estimate of total fertility, \( T \).
Applying the same $\alpha$ and $\beta$ to the standard gompits for the ages that divide conventional age groups (i.e. 20, 25 … 50), applying the anti-gompit transformation, and multiplying by $T^\hat{\gamma}$ produces a scaled cumulated fertility schedule. Differencing successive estimates of cumulated fertility and dividing by five produces the fitted fertility schedule for conventional age groups (15-19; 20-24 etc.) even if the data were initially classified with a half-year shift. (If the model has been fitted using only the $F$-points, then $\alpha$ and $\beta$ are defined by the $F$-line only. The smoothed fertility schedule is produced by a series of steps identical to that described above except that the fitted proportions are multiplied by the level of fertility estimated from the recent data themselves, rather than by an estimate based on the parity data.)

3.3.2 Reverse survival model
Reverse survival is a method for estimating fertility from data collected in a census or single-round survey that can be used even if no questions have been asked about fertility directly. In a population closed to migration, the population of any age $x$ are the survivors of the births in that population $x$ completed years previously. The application of the concept of reverse survival described here produces annual estimates of the General Fertility and Total Fertility Ratios for up to 15 years before the inquiry. To calculate the General Fertility Ratio requires only that one estimates the past size of the population of adult women from the number of women enumerated in an inquiry by allowing for adult mortality. To calculate Total Fertility, however, requires information not only on the number of births occurring each year, but also on the ages of the mothers of these newborn children. One relatively simple way of estimating this information, if it is not known, is to apportion births to age groups of mother using independent estimates of the age distribution of fertility.

3.3.2.1 Data required
To derive General Fertility Ratios for individual years, the following data are required:

- Tabulations of the population (of both sexes) aged 0 to 14, by single years of age.
- Tabulations of the female population aged 15 to 64 by five-year age group.
- Cohort survival probabilities, $L_x$, for children aged 0 to 14 of both sexes.
- Survivorship ratios, $5L_x-5/5L_x$ for adult women for each of the three five-year periods preceding the inquiry.
In respect of the mortality estimates, the implementation of the method in the associated Excel workbook allows these to be specified either by reference to period-specific parameters $\alpha$ and $\beta$ of appropriate relational model life tables, or to identified values of $5q_0$ (for children) and $45q_{15}$ (for adult women) for each of the three five-year periods preceding the enquiry.

To produce estimates of Total Fertility, one also requires either

- A single age-specific fertility distribution that is assumed to apply to the entire period covered by the estimates, or
- Two age-specific fertility distributions, one of which applies to a date reasonably close to the index inquiry and the other to a date approximately 15 years prior to that.

Either a series of fertility rates or the parameters of a relational Gompertz model fitted to a standard fertility schedule can be used as an input to the calculations. Note that only the estimated shape of the fertility distribution is based on these fertility schedules. It is the estimated number of births relative to the population of women of childbearing age that almost entirely determines the estimates of the General and Total Fertility Ratios.

### 3.3.2.2 Assumptions

The population is assumed to have been closed to migration for as many years as are covered by the reverse survival estimates. However, because children usually migrate with their mothers, errors in the numerator and denominator of the estimated rates largely cancel out. Significant bias will result only if migration flows are large and migrants have different fertility from the rest of the population.

### 3.3.2.3 Application of the method

**Step 1: Estimate the number of births in each year before the inquiry**

The enumerated population aged $x$ in any inquiry represents the survivors from the births that occurred in the 12-month period centered on the date $x+0.5$ years before the inquiry. Algebraically, $B_{x+0.5} = N_x c L_x \square, 0 \leq x \leq 14$.

The measure of survivorship used in this calculation, $L_x$, is a cohort survival factor. It depends on mortality at successive ages in successive years leading up to the inquiry. Appropriate cohort
estimates of mortality may be available from the inquiry used to estimate fertility. Such estimates include indirect estimates from children ever-born and surviving data, in the case of a census, and direct estimates from a cohort analysis of the birth histories, in the case of a fertility survey. Brass (1979) describes a simple procedure for estimating $L_x$ directly from proportions of children surviving according to their mothers’ age group. If a series of cohort survival ratios is not readily to hand, one can be derived from estimates of period mortality by single years of age for each five-year period before the inquiry. Working with a relational logit system of model life tables, define $\alpha_T$ and $\beta_T$ to be the parameters that generate a life table for period $T$, where $T=0$ refers to the period 0-4 years before the inquiry, $T=5$ to the period 5-9 years before the inquiry and $T=10$ to the period 10-14 years before the inquiry. For a given standard, indexed by the superscript $s$, $Y_x, T=\alpha_T+\beta_T.Y_{sx}$ where $Y$ is the logit function: 

$$Y_x=\frac{1}{2}\ln\left(\frac{1-l_x}{l_x}\right)$$

Equation 1 and $l(x)$ or $(p_0)$ refers to the proportion surviving from birth to exact age $x$ in a life table with a radix of 1. Equation 1 implies that $l_x=(1+\exp(2Y_x))-1$ And that

$$l_x,T=(1+\exp(2(\alpha_T+\beta_T.Y_{sx})))^{-1}$$

Equation 2

At ages other than infancy, one can approximate $L_{x,T}$, the person-years lived between $x$ and $x+1$ in period $T$ by assuming that survivorship declines linearly on the logistic scale and, therefore, that the logit of $L_{x,T}$ is the average of $Y_{x,T}$ and $Y_{x+1,T}$. From Equation 2 

$$L_x,T=l_x+0.5=(1+\exp(2(Y_x,T+Y_{x+1},T2)))-1=(1+\exp(Y_x,T+Y_{x+1},T))^{-1}$$

Equation 3

The values of $l(x)$ in systems of model life tables are often tabulated by single ages up to age 5, and then at every fifth year of age. If this is the case for children aged 5 or more, one can assume that logit survivorship declines linearly over the entire age range $x$ to $x+5$. Thus, for example, if a life table has tabulated values at $x=5$ and $x=10$, the estimated value of $L_{9,T}$ would be given by $L_{9,T}=l_{9.5,T}=(1+\exp(2(110Y_{5,T}+910Y_{10,T})))^{-1}=(1+\exp(0.2Y_{5,T}+1.8Y_{10,T}))-1$

Equation 4

For infants, one should allow for the concentration of deaths in the first days and weeks of life. In medium and high mortality populations, one can approximate person years lived in the first year of life as $L_{0,T}=0.3+0.7(\exp(2Y_{1,T}))-1$

Equation 4

Survivorship ratios from one age to the next, $P_{x,T}$, in time period $T$, are derived from the ratio of successive values of $L_{x,T}$: 

$$P_{x,T}=L_x,TL_x,0\leq x\leq 14 P_{0,T}=L0,T$$

Equation 5
Once estimates of the survivorship of children by single years of age and five-year time periods have been obtained in either the way just outlined or by some other procedure, an estimate of cohort survivorship by single years of age for single-year age cohorts can be calculated as follows. Recall that \( P_{a,T} \) is the survivorship ratio between ages \( a \) and \( a+1 \) in time period \( T \) (where \( T=0, 5 \) or 10, corresponding to five-year periods 0-4, 5-9 and 10-14 years before the inquiry). Further define \( S_{a,t} \) to be the survivorship ratio between ages \( a \) and \( a+1 \) in the period \( t \) to \( t+1 \) years before the inquiry, \( 0 \leq t \leq 14 \). Using linear interpolation to estimate survivorship for the intermediate years:

\[
S_{a,t} = \begin{cases} 
P_{a,0} & 0 \leq t \leq 2 \\
P_{a,T}(1-t-T) + P_{a,T}(t-T) & 2 < t < 8 \\
P_{a,10} & 12 < t < 15 
\end{cases}
\]

The proportion of births occurring \( x \) to \( x+1 \) years earlier that survive to the time of the inquiry, \( cL_x \), can then be calculated as

\[
L_x = S_{0,x}S_{1,x-1}...S_{x-1,1}S_x
\]

The number of births in each year before an inquiry, centered on the point mid-way through that year (i.e. 6 months before the date of the inquiry), is thus

\[
B_{x+0.5} = N_x cL_x, 0 \leq x \leq 14
\]

where \( N_x \) is the number of children aged \( x \) reported in the inquiry.

**Step 2: Estimate the mid-year populations of women by five-year age group**

The calculation of survivorship for women aged 15 to 64 at the date of the inquiry can be done in a straightforward way because mortality is usually fairly low in the central adult ages. Thus, even approximate estimates of mortality will enable one to produce a satisfactory estimate of the past population from the enumerated population. The absolute variation of mortality with age within any five-year age group is small. Therefore, one can therefore approximate \( sL_x \) by linear interpolation between \( Y_x \) and \( Y_{x+5} \). This means that one can estimate survivorship between five-year age groups at time \( T \) (where \( T=5, 10 \) and 15) as

\[
P_{5x,T} = L_{5x} + 5L_{5x} = 1 + \exp(2aT + \beta T(Y_{5x} + Y_{5x+5}))1 + \exp(2aT + \beta T(Y_{5x+5} + Y_{5x+10}))
\]

Starting with the population enumerated at \( T = 0 \), the number of women in each five-year age group \( T+5 \) years before the inquiry can then be calculated from the number at \( T \):

\[
N_{5x,T+5} = N_{5x+5,T+5}, 10 \leq x < 60, T = 0, 5, 10
\]

Furthermore, since the age structure of a population changes only slowly, the mid-year populations of women in the age groups 10-14, 15-19, …, 60-64 for each year before the date of the inquiry can be estimated by interpolating linearly between the population estimates for 0, 5,
10 and 15 years before the inquiry produced by Equation 8. For example, to estimate the number of women aged 20-24 at the point 8½ years before the inquiry, the formula would be

\[ 5N_{20,8.5} = 0.3(5N_{20,5}) + 0.7(5N_{20,10}) \]

**Step 3: Derive General Fertility Ratios**

The General Fertility Ratio for the year centered on the point \( x-0.5 \) years before the inquiry is

\[ GFR_x + 0.5 = B_x + 0.5 \sum_{a=15}^{5455} N_{x,a} + 0.5 \sum_{0\leq x\leq 14} \]

Equation 9

where the denominator is the total number of women aged between 15 and 49 at the mid-point of the year during which the births occurred.

**Step 4: Estimate age-specific and Total Fertility**

A natural extension to the calculation of General Fertility Ratios is to make use of a schedule describing the age-pattern of fertility in the population being studied to estimate Total Fertility by a procedure akin to indirect standardization.

Such a fertility schedule might come from data on recent births collected in the same inquiry as is being analyzed by reverse survival methods. As only information on the age pattern of fertility is obtained from these data, no need exists to adjust the reports for reference-period errors. If a second fertility schedule is available from a previous census or survey of the same population (preferably conducted about 15 years earlier) one can interpolate between the two schedules, or if necessary extrapolate, to estimate the shape of the fertility schedule in each year for which one intends to estimate Total Fertility. However, even if one suspects that fertility has been changing, having two schedules is not essential, since the age pattern of fertility evolves only gradually and estimates of Total Fertility are not very sensitive to the exact assumptions made about the shape of the fertility distribution. Thus, a single schedule (preferably referring to about the middle of the estimation period) will suffice.

If fertility has been estimated using a relational Gompertz model, in combination with the chosen standard, the parameters \( \alpha \) and \( \beta \) derived during the estimation of fertility define the shape of the fertility schedule. Accordingly, values of \( \alpha \) and \( \beta \) can be used to determine the shape of the fertility schedule(s) used in the estimation of Total Fertility. Once one has estimated the proportion of Total Fertility occurring in each age group can for each year before the inquiry, these proportions can be applied to the population of women in each age group in each year to
estimate the number of births that would have occurred to women in that age group if Total Fertility equaled one child per woman. Thus, once one has selected a fertility schedule (scaled to a Total Fertility of 1), $5f_{a,x+0.5}$, for each age group ($a=15, 20, \ldots, 45$) for each of the 15 years ($x$) before an inquiry, the expected number of births to women in each age group in each year is given by $5B^*a,x+0.5=5Na,x+0.5.5f^a,x+0.5,0 \leq x \leq 14$

It follows that the total number of births that would have occurred in year $x$ if Total Fertility had equaled 1 is $B^*x+0.5=\sum a=15,5455Na,x+0.5.5f^a,x+0.5,0 \leq x \leq 14$. However, Step 2 yielded an estimate of the actual number of births in each year, $B_x+0.5$. The estimate of Total Fertility for each year is thus the ratio of $B$ to $B^*TFx+0.5=Bx+0.5B^*x+0.5,0 \leq x \leq 14$. Estimates of the age-specific fertility rates for the year can be produced by multiplying the proportional rates, $5f^a,x$, for each age group ($a$) and year ($x$) by the estimate of Total Fertility for that year. It should be noted, however, that the results will merely reproduce the age pattern of fertility that was input into the calculations: this method provides no new information on the age pattern of fertility.
CHAPTER FOUR
LEVELS AND TRENDS OF FERTILITY IN KENYA

4.1 Introduction
This chapter discusses data quality issues with 2009 census data (age and parity misreporting), results from national fertility, fertility by province, residence (Rural and urban) and education level.

4.2 Data quality
Census data is often characterized by content errors arising from age and parity misreporting (CBS, 2002; GSS, 2013). The assessment of data quality was done using graphs, calculation of Whipples and Myers blended indices and assessment of the proportion childless.

4.2.1 Visual inspection of graphs and tabulations of reported age
The age distribution of a population is determined by fertility, mortality and migration. It, therefore, follows a fairly predictable pattern. The “expected” pattern of numbers at various ages is that, given stable birth rates, there should be more people in an age compared to the next higher age as a result of death in the absence of migration. Thus, for example, there should be fewer people aged 10 years than aged 9 years since not all nine-year olds will survive to age ten (GSS, 2013). By visually inspecting data on reported age, we can identify errors arising from digit preference. The 2009 population and housing census revealed that there was preference for ages ending with 0, 5, 2 and 8. For instance, the number of females aged 19 years was 356,175 and those aged 20 years were 516,549, a difference of 160,374. This difference can be attributed to the preference of digit 20 compared to 19. Similarly, those who reported to be 34 years were 189,997 and those aged 35 years were 316,609, a difference of 126,612 which can also be attributed to preference of digit 5.

Graphing data is another visual approach for identifying age misreporting. Figure 4.1 shows the distribution of the population by single years for males and females. The graph shows preference for digit 0 and 5 as portrayed by the observed peaks followed by even numbers and the avoidance of odd numbers shown by troughs for 1, 3, 7 and 9.
Figure 4.1 Age distribution in single years (2009 census)
4.2.2 Measuring age misreporting

In an attempt to quantify the extent of age misreporting, further analysis of age misreporting was done using Whipples and Myers blended indices. The Whipples index is used to detect avoidance or prevalence of a particular digit or each terminal digit. The values of the index vary from 100 to 500. A data set is said to be highly inaccurate if the index is above 175, inaccurate if the index is between 125 and 175, fairly acceptable if it is between 110 and 125, and highly accurate if it is less than 105. The value calculated from the 2009 census data was 148, implying that the data was inaccurate and needed adjustment before use. Further, a Myers index of 15 for females was arrived at which meant preference of digit 0 and 5. Figure 4.2 shows the ages that were avoided and preferred. Among the digits that were highly avoided are, 1, 3 and 9 while 0 and 5 were preferred. This problem of age misreporting was corrected by grouping the population in 5 year age groups. Although grouping the data in five year age groups smoothen it, figure 4.3 shows that there is a peak in age group 20-24 due to a possible misclassification of ages from those aged 15-19 hence further refinement was done by adjusting the population in the five year age groups by the five point method (smoothening). Figure 4.3 and 4.4 shows the distribution of female population by five year age groups for reported and adjusted ages respectively. When the data was adjusted, the curve is smoother compared to when not especially for the age group 20-24 as revealed in figure 4.4.

Figure 4.2 Myers index for digit preference.
Figure 4.3 Distribution of reported age by 5 year age groups.
4.2.3 Parity misreporting

The analysis of reported parities of the 2009 census had a problem of childless women and women not stating their parity. When the proportion for unknown parity is less than 2% in each age group, there is no need of adjustment. This proportion ($U_i$) was more than 2% in three age groups (15-19, 20-24 and 25-29) as revealed by Table 4.1. Further analysis was done by plotting the proportion of women whose parity was not stated against the proportion childless as shown in figure 4.5. The points were found to lie approximately on a straight line which is evidence of misclassification. This problem was solved by the application of the El badry (1961) method after which the points were now not approximately lying on a straight line hence ready for use.
Table 4.1 Reported parities and proportion for unknown parity ($u_i$)

<table>
<thead>
<tr>
<th>Age group</th>
<th>15 - 19</th>
<th>20 - 24</th>
<th>25 - 29</th>
<th>30 - 34</th>
<th>35 - 39</th>
<th>40 - 44</th>
<th>45 - 49</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parity 0</td>
<td>217,216</td>
<td>237,994</td>
<td>122,164</td>
<td>44,645</td>
<td>22,393</td>
<td>13,683</td>
<td>11,471</td>
</tr>
<tr>
<td>Parity 1</td>
<td>367,799</td>
<td>587,588</td>
<td>330,427</td>
<td>128,420</td>
<td>59,990</td>
<td>32,164</td>
<td>23,110</td>
</tr>
<tr>
<td>Parity 2</td>
<td>209,905</td>
<td>449,905</td>
<td>393,927</td>
<td>227,382</td>
<td>122,607</td>
<td>63,865</td>
<td>40,380</td>
</tr>
<tr>
<td>Parity 3</td>
<td>168,992</td>
<td>252,324</td>
<td>310,349</td>
<td>227,422</td>
<td>154,733</td>
<td>90,740</td>
<td>60,791</td>
</tr>
<tr>
<td>Parity 4</td>
<td>178,155</td>
<td>143,316</td>
<td>222,726</td>
<td>201,435</td>
<td>151,862</td>
<td>100,295</td>
<td>77,751</td>
</tr>
<tr>
<td>Parity 5</td>
<td>161,321</td>
<td>81,253</td>
<td>123,611</td>
<td>155,112</td>
<td>131,680</td>
<td>91,442</td>
<td>78,986</td>
</tr>
<tr>
<td>Parity 6</td>
<td>166,591</td>
<td>77,909</td>
<td>67,713</td>
<td>114,760</td>
<td>116,292</td>
<td>85,355</td>
<td>77,839</td>
</tr>
<tr>
<td>Parity 7</td>
<td>149,089</td>
<td>43,811</td>
<td>34,368</td>
<td>68,732</td>
<td>88,655</td>
<td>73,849</td>
<td>69,675</td>
</tr>
<tr>
<td>Parity 8</td>
<td>146,769</td>
<td>45,462</td>
<td>33,941</td>
<td>44,803</td>
<td>70,028</td>
<td>67,760</td>
<td>65,001</td>
</tr>
<tr>
<td>Parity 9</td>
<td>96,872</td>
<td>31,686</td>
<td>8,394</td>
<td>20,121</td>
<td>37,765</td>
<td>43,653</td>
<td>46,015</td>
</tr>
<tr>
<td>Parity 10+</td>
<td>181,497</td>
<td>62,427</td>
<td>18,603</td>
<td>25,965</td>
<td>45,414</td>
<td>68,762</td>
<td>85,836</td>
</tr>
<tr>
<td>Not stated(U)</td>
<td>162,829</td>
<td>78,136</td>
<td>44,094</td>
<td>28,482</td>
<td>18,816</td>
<td>13,171</td>
<td>10,191</td>
</tr>
<tr>
<td>Total</td>
<td>2,207,035</td>
<td>2,091,811</td>
<td>1,710,317</td>
<td>1,287,279</td>
<td>1,020,235</td>
<td>744,739</td>
<td>647,046</td>
</tr>
</tbody>
</table>

$U_i$:
- 0.07
- 0.04
- 0.03
- 0.02
- 0.02
- 0.02
- 0.02

$Z_i$:
- 0.10
- 0.11
- 0.07
- 0.03
- 0.02
- 0.02
- 0.02
Table 4.2 Corrected parities and proportion for unknown parity (ui)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>360383</td>
<td>297494</td>
<td>151021</td>
<td>61659</td>
<td>32120</td>
<td>20219</td>
<td>15898</td>
</tr>
<tr>
<td>1</td>
<td>367799</td>
<td>587588</td>
<td>330427</td>
<td>128420</td>
<td>59990</td>
<td>32164</td>
<td>23110</td>
</tr>
<tr>
<td>2</td>
<td>209905</td>
<td>449905</td>
<td>393927</td>
<td>227382</td>
<td>122607</td>
<td>63865</td>
<td>40380</td>
</tr>
<tr>
<td>3</td>
<td>168992</td>
<td>252324</td>
<td>310349</td>
<td>227422</td>
<td>154733</td>
<td>90740</td>
<td>60791</td>
</tr>
<tr>
<td>4</td>
<td>178155</td>
<td>143316</td>
<td>222726</td>
<td>201435</td>
<td>151862</td>
<td>100295</td>
<td>77751</td>
</tr>
<tr>
<td>5</td>
<td>161321</td>
<td>81253</td>
<td>123611</td>
<td>155112</td>
<td>131680</td>
<td>91442</td>
<td>78986</td>
</tr>
<tr>
<td>6</td>
<td>77909</td>
<td>67713</td>
<td>114760</td>
<td>116292</td>
<td>85355</td>
<td>77839</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>43811</td>
<td>34368</td>
<td>68732</td>
<td>88655</td>
<td>73849</td>
<td>69675</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>45462</td>
<td>33941</td>
<td>44803</td>
<td>70028</td>
<td>67760</td>
<td>65001</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>8394</td>
<td>20121</td>
<td>37765</td>
<td>43653</td>
<td>46015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>18603</td>
<td>25965</td>
<td>45414</td>
<td>68762</td>
<td>85836</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>19662</td>
<td>18636</td>
<td>15237</td>
<td>11468</td>
<td>9089</td>
<td>6635</td>
<td>5764</td>
</tr>
<tr>
<td>TOTAL</td>
<td>1466217</td>
<td>1997698</td>
<td>1710317</td>
<td>1287279</td>
<td>1020235</td>
<td>744739</td>
<td>647046</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>UI</th>
<th>Zi</th>
<th>Ui</th>
<th>Zi</th>
<th>Ui</th>
<th>Zi</th>
<th>Ui</th>
<th>Zi</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>0.15</td>
<td>0.09</td>
<td>0.05</td>
<td>0.03</td>
<td>0.03</td>
<td>0.02</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2 shows that the proportion for unknown parity is now less than 2% in every age group implying that the quality of the data has been improved hence ready for use.

In summary, the values of Whipples and Myers indices were 148 and 15 respectively implying that the data collected in 2009 Kenyan census was inaccurate hence there was need for adjustment in order to improve it’s quality before use as detailed in this chapter.
4.3 Results of TFR estimates

This section presents results of TFR by the relational gompertz and reverse survival models as described in chapter three of this project. First are results nationally then by selected background characteristics of women of reproductive ages.

4.3.1 National estimates of TFR by relational and reverse survival models

The average parities show a general increase from 0.186 for 15-19 age group to a high of 4.8 for 45-49 age group as revealed by Table 4.3. The results from the relational gompertz model show that Total fertility calculated from lifetime fertility was 4.83 while that calculated from observed ASFRs was 4.42 as shown by table 4.3. The reported TFR was lower (4.42) than the estimated TFR by the Relational gompertz model, an indication of underreporting of births in the 12 months preceding the census or overreporting of lifetime fertility. The observed discrepancy in TFR by the relational gompertz model could also be attributed to displacement of births which
happens when dates of childbirths are transferred by the enumerators. This is common in Sub-Saharan Africa (Pullum 2006). This problem often leads to underestimation or overestimation of fertility. Estimates of ASFRs decreased sharply in all the age groups except for the 20-24 age group in both 1999 and 2009 censuses.

Reported fertility experience as reported in the census, has been found to be distorted for a variety of reasons, including recall lapse (GSS, 2013). This was evident in the Kenya 2009 census. The P/F ratios which relates the age pattern of fertility derived from information on recent births to the level of fertility implied by the average parity of women in age groups 20-24, 25-29 and perhaps 30-34 years, are expected to be one if the expected mean parity (Pi) is equal to the cumulative age specific fertility (Fi). A P/F ratio greater than one implies underreporting of births in the last 12 months and lower than one denotes that parity might have been underreported. The P/F ratios for 2009 were greater than one, implying that current births were under-reported. This was more pronounced in the 15-19 years age group in 2009 census (Table 4.3). Disregarding the P/F ratio for the 15-19 years age group because of underreporting of births and high infant mortality, the P/F ratios suggest that current births were underreported by about 11 percent to 27 percent (Table 4.4).

Having presented results by relational model, this study attempted to estimate fertility level and trend using the reverse survival method. To do this, a simulated population was first projected over 15 years using a set of fertility and mortality age and sex patterns. The projected population was then reverse survived using the Excel template FE_reverse_4.xlsx, provided with Timæus and Moultrie (2012). The results reveals that fertility in Kenya has been reducing from a high of 5.6 in 1995 to 4.3 in 2009 children per woman (Table 4.4). However, the results also revealed that fertility in Kenya had stalled at 4.8 children in 2000 to 2007 (Figure 4.6). The high peak and trough observed in 1999 and 1998 respectively were smoothed by three point moving average as shown in figure 4.6. The General fertility rate varied from a high of 207 in 1999 and a low of 137 births per thousand women (Table 4.4).
Table 4.3  Distribution of average parities, age specific fertility rates and Total fertility nationally

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Observed ASFR</th>
<th>Observed parities</th>
<th>Fitted ASFRS(2009)</th>
<th>Fitted MCEB</th>
<th>P/F Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 - 19</td>
<td>0.0730</td>
<td>0.2123</td>
<td>0.1124</td>
<td>0.1860</td>
<td></td>
</tr>
<tr>
<td>20 - 24</td>
<td>0.2161</td>
<td>1.2638</td>
<td>0.2448</td>
<td>1.1616</td>
<td>1.2719</td>
</tr>
<tr>
<td>25 - 29</td>
<td>0.2143</td>
<td>2.4202</td>
<td>0.2392</td>
<td>2.4068</td>
<td>1.1741</td>
</tr>
<tr>
<td>30 - 34</td>
<td>0.1753</td>
<td>3.5185</td>
<td>0.1876</td>
<td>3.4847</td>
<td>1.1699</td>
</tr>
<tr>
<td>35 - 39</td>
<td>0.1233</td>
<td>4.2748</td>
<td>0.1254</td>
<td>4.2697</td>
<td>1.1480</td>
</tr>
<tr>
<td>40 - 44</td>
<td>0.0601</td>
<td>4.6388</td>
<td>0.0518</td>
<td>4.7163</td>
<td>1.1161</td>
</tr>
<tr>
<td>45 - 49</td>
<td>0.0219</td>
<td>4.8897</td>
<td>0.0057</td>
<td>4.8337</td>
<td>1.1461</td>
</tr>
<tr>
<td>TFR</td>
<td>4.4203</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4 Fertility estimates per year(1995-2009)

<table>
<thead>
<tr>
<th>Mid-Year</th>
<th>GFR (15-49)</th>
<th>Total Fertility</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009.144</td>
<td>137</td>
<td>4.3</td>
</tr>
<tr>
<td>2008.144</td>
<td>130</td>
<td>4.0</td>
</tr>
<tr>
<td>2007.144</td>
<td>159</td>
<td>4.8</td>
</tr>
<tr>
<td>2006.144</td>
<td>159</td>
<td>4.8</td>
</tr>
<tr>
<td>2005.144</td>
<td>167</td>
<td>4.9</td>
</tr>
<tr>
<td>2004.144</td>
<td>168</td>
<td>4.9</td>
</tr>
<tr>
<td>2003.144</td>
<td>173</td>
<td>4.9</td>
</tr>
<tr>
<td>2002.144</td>
<td>165</td>
<td>4.6</td>
</tr>
<tr>
<td>2001.144</td>
<td>178</td>
<td>5.0</td>
</tr>
<tr>
<td>2000.144</td>
<td>178</td>
<td>4.9</td>
</tr>
<tr>
<td>1999.144</td>
<td>207</td>
<td>5.6</td>
</tr>
<tr>
<td>1998.144</td>
<td>151</td>
<td>4.0</td>
</tr>
<tr>
<td>1997.144</td>
<td>204</td>
<td>5.4</td>
</tr>
<tr>
<td>1996.144</td>
<td>186</td>
<td>4.8</td>
</tr>
<tr>
<td>1995.144</td>
<td>187</td>
<td>4.8</td>
</tr>
</tbody>
</table>
4.3.2 Differentials of fertility by region  (Relational gompertz model)
Estimates of fertility by provinces reveal that, TFR decreased in all the provinces in 2009 census compared to the levels in 1999 except Northeastern province where there was an increase of 1.5 births. This could be attributed to the inflation of data collected in several northeastern districts (KNBS, 2010) as revealed in Figure 4.7. Nairobi province had the least TFR of 3.3 followed by Central province. Central province experienced a significant drop in TFR of 1.5 births while Western, Nyanza and Coast provinces having the least drop in TFR as revealed in figure 4.7.
4.3.3 Differentials of fertility by place of residence (Relational gompertz model -2009)

Table 4.7 and 4.8 reveals that in 2009, fertility in Rural (TFR of 6.045) areas is higher than that in urban (TFR of 4.125) areas, a difference of more than 2 births. As expected, the age specific fertility rates decreased in all age groups except for 20-24 and 25-29 age groups. The difference between the TFR calculated from reported ASFRs and fitted ASFRs is higher in rural than in urban indicating that under reporting of births is higher in rural than in urban. The P/F ratios in rural areas were all greater than one implying current births were underreported. This was more pronounced in age group 45-49 among women living in rural areas. Table 4.5 reveals that underreporting of current births was even higher for women in the age group 45-49 compared to their rural counterparts, underreporting of current births was 34% as revealed in table 4.6.
Table 4.5 Fertility in rural Kenya by relational gompertz model

<table>
<thead>
<tr>
<th>Age group</th>
<th>Observed ASFRS</th>
<th>Observed MCEB</th>
<th>Fitted ASFRS</th>
<th>Fitted MCEB</th>
<th>P/F Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-19</td>
<td>0.0780</td>
<td>0.2255</td>
<td>0.1430</td>
<td>0.2660</td>
<td></td>
</tr>
<tr>
<td>20-24</td>
<td>0.2490</td>
<td>1.4807</td>
<td>0.2820</td>
<td>1.4120</td>
<td>1.2720</td>
</tr>
<tr>
<td>25-29</td>
<td>0.2450</td>
<td>2.8528</td>
<td>0.2830</td>
<td>2.8560</td>
<td>1.1850</td>
</tr>
<tr>
<td>30-34</td>
<td>0.1960</td>
<td>4.1575</td>
<td>0.2360</td>
<td>4.1670</td>
<td>1.1850</td>
</tr>
<tr>
<td>35-39</td>
<td>0.1400</td>
<td>5.1502</td>
<td>0.1730</td>
<td>5.1960</td>
<td>1.1850</td>
</tr>
<tr>
<td>40-44</td>
<td>0.0690</td>
<td>5.9171</td>
<td>0.0810</td>
<td>5.8450</td>
<td>1.2180</td>
</tr>
<tr>
<td>45-49</td>
<td>0.0250</td>
<td>6.4190</td>
<td>0.0110</td>
<td>6.0430</td>
<td>1.2870</td>
</tr>
<tr>
<td>TFR</td>
<td>5.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.6 Fertility in urban Kenya by relational gompertz model

<table>
<thead>
<tr>
<th>Age group</th>
<th>Observed ASFrs</th>
<th>Observed MCEB</th>
<th>Fitted ASFrs</th>
<th>Fitted MCEB</th>
<th>P/F Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-19</td>
<td>0.0610</td>
<td>0.1794</td>
<td>0.0800</td>
<td>0.1280</td>
<td></td>
</tr>
<tr>
<td>20-24</td>
<td>0.1700</td>
<td>0.9381</td>
<td>0.1920</td>
<td>0.8590</td>
<td>1.1160</td>
</tr>
<tr>
<td>25-29</td>
<td>0.1730</td>
<td>1.7940</td>
<td>0.2020</td>
<td>1.8750</td>
<td>1.1020</td>
</tr>
<tr>
<td>30-34</td>
<td>0.1410</td>
<td>2.7601</td>
<td>0.1690</td>
<td>2.8180</td>
<td>1.0590</td>
</tr>
<tr>
<td>35-39</td>
<td>0.0920</td>
<td>3.5668</td>
<td>0.1210</td>
<td>3.5490</td>
<td>1.1180</td>
</tr>
<tr>
<td>40-44</td>
<td>0.0400</td>
<td>4.2296</td>
<td>0.0540</td>
<td>3.9930</td>
<td>1.2030</td>
</tr>
<tr>
<td>45-49</td>
<td>0.0150</td>
<td>4.8230</td>
<td>0.0070</td>
<td>4.1200</td>
<td>1.3440</td>
</tr>
<tr>
<td>TFR</td>
<td>3.46</td>
<td></td>
<td></td>
<td></td>
<td>4.125</td>
</tr>
</tbody>
</table>

4.3.4 Differentials of fertility by level of education (Relational gompertz model -2009)

The 2009 census collected information on highest level of education attended by women aged 15-49 years. Tables 4.7, 4.8 and 4.9 displays reported and fitted fertility rates by education attainment of females aged 15-49 years in 2009. TFR is inversely related to the level of education as established in Kenya (figure 4.8) and other African countries (Central Statistical Office and Macro International Inc., 2008; KNBS and ICF Macro, 2011; Liberia Institute of Statistics and Geo-Information Service and Ministry of Health and Social Welfare, 2008; Zimbabwe National Statistics Agency and ICF International, 2012). In 2009, TFR varied from 6.95 among women who had never attended school to 3.88 among females with secondary and above education. The level of fertility among women who had never attended school was 1.8 times (almost twice) that of females with secondary and above education. With regard to P/F ratios, there was a general decline of P/F ratios in all categories of level of education attainment. Women who had never attended school in the age group 20-24 had the highest percentage of
underreporting of current births (59%) while women with secondary and above education in the age group 25-29 had the least percentage of underreporting (1%).

Table 4.7 Fertility level of Never attended

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Observed ASFR</th>
<th>Observed parities</th>
<th>Fitted ASFRS</th>
<th>Fitted MCEB</th>
<th>P/F Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 - 19</td>
<td>0.0916</td>
<td>0.3773</td>
<td>0.1639</td>
<td>0.2923</td>
<td></td>
</tr>
<tr>
<td>20 - 24</td>
<td>0.2400</td>
<td>1.8139</td>
<td>0.3355</td>
<td>1.6466</td>
<td>1.5852</td>
</tr>
<tr>
<td>25 - 29</td>
<td>0.2479</td>
<td>3.3756</td>
<td>0.3326</td>
<td>3.3578</td>
<td>1.4318</td>
</tr>
<tr>
<td>30 - 34</td>
<td>0.2052</td>
<td>4.9719</td>
<td>0.2709</td>
<td>4.8817</td>
<td>1.4355</td>
</tr>
<tr>
<td>35 - 39</td>
<td>0.1550</td>
<td>6.0137</td>
<td>0.1911</td>
<td>6.0428</td>
<td>1.3873</td>
</tr>
<tr>
<td>40 - 44</td>
<td>0.0794</td>
<td>6.6182</td>
<td>0.0854</td>
<td>6.7462</td>
<td>1.3531</td>
</tr>
<tr>
<td>45 - 49</td>
<td>0.0342</td>
<td>6.9583</td>
<td>0.0106</td>
<td>6.9487</td>
<td>1.3794</td>
</tr>
<tr>
<td>TFR</td>
<td>5.2657</td>
<td>6.9501</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.8 Fertility level of Primary

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Observed ASFR</th>
<th>Observed parities</th>
<th>Fitted ASFRS</th>
<th>Fitted MCEB</th>
<th>P/F Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 - 19</td>
<td>0.0898</td>
<td>0.2424</td>
<td>0.1608</td>
<td>0.3441</td>
<td></td>
</tr>
<tr>
<td>20 - 24</td>
<td>0.2707</td>
<td>1.6424</td>
<td>0.2701</td>
<td>1.5008</td>
<td>1.2202</td>
</tr>
<tr>
<td>25 - 29</td>
<td>0.2353</td>
<td>2.8171</td>
<td>0.2594</td>
<td>2.8423</td>
<td>1.0877</td>
</tr>
<tr>
<td>30 - 34</td>
<td>0.1864</td>
<td>3.9877</td>
<td>0.2166</td>
<td>4.0397</td>
<td>1.0910</td>
</tr>
<tr>
<td>35 - 39</td>
<td>0.1331</td>
<td>4.9574</td>
<td>0.1623</td>
<td>4.9913</td>
<td>1.1122</td>
</tr>
<tr>
<td>40 - 44</td>
<td>0.0663</td>
<td>5.7744</td>
<td>0.0804</td>
<td>5.6142</td>
<td>1.1665</td>
</tr>
<tr>
<td>45 - 49</td>
<td>0.0216</td>
<td>6.3225</td>
<td>0.0119</td>
<td>5.8175</td>
<td>1.2444</td>
</tr>
<tr>
<td>TFR</td>
<td>5.0163</td>
<td>5.8076</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.9 Fertility level of Secondary and above

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Observed ASFR</th>
<th>Observed parities</th>
<th>Fitted ASFRS</th>
<th>Fitted MCEB</th>
<th>P/F Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 - 19</td>
<td>0.0342</td>
<td>0.0946</td>
<td>0.0461</td>
<td>0.0498</td>
<td></td>
</tr>
<tr>
<td>20 - 24</td>
<td>0.1453</td>
<td>0.6345</td>
<td>0.1707</td>
<td>0.6076</td>
<td>1.0617</td>
</tr>
<tr>
<td>25 - 29</td>
<td>0.1764</td>
<td>1.5301</td>
<td>0.2072</td>
<td>1.5996</td>
<td>1.0061</td>
</tr>
<tr>
<td>30 - 34</td>
<td>0.1473</td>
<td>2.5425</td>
<td>0.1770</td>
<td>2.5835</td>
<td>1.0746</td>
</tr>
<tr>
<td>35 - 39</td>
<td>0.0924</td>
<td>3.3542</td>
<td>0.1207</td>
<td>3.3363</td>
<td>1.1316</td>
</tr>
<tr>
<td>40 - 44</td>
<td>0.0377</td>
<td>3.9659</td>
<td>0.0484</td>
<td>3.7630</td>
<td>1.2101</td>
</tr>
<tr>
<td>45 - 49</td>
<td>0.0113</td>
<td>4.4827</td>
<td>0.0049</td>
<td>3.8696</td>
<td>1.3422</td>
</tr>
<tr>
<td>TFR</td>
<td>3.2225</td>
<td>3.8751</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The pattern of fertility shown by reported TFRs is also reflected by the mean number of children ever born to women aged 45-49 years. Regarding reported TFRs, rural women and women who had never attended school reported the highest mean number of children ever born 6.4 and 6.9 respectively.

Table 4.10 TFR estimates by various models

<table>
<thead>
<tr>
<th>Method</th>
<th>TFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reverse survival model</td>
<td>4.34</td>
</tr>
<tr>
<td>Relational Gompertz model</td>
<td>4.83</td>
</tr>
<tr>
<td>Brass consistency check</td>
<td>4.87</td>
</tr>
</tbody>
</table>

Table 4.10 shows results of the three models. The results reveals that although different methods were used to estimate Kenyan fertility at national with different assumptions, the results obtained were very close ranging from a TFR of 4.34 to 4.83.

4.4 Discussion

The aim of this study was to derive fertility estimates for Kenya using reverse sevival and the relational gompertz models. These estimates were compared with that from Brass consistency check (Table 4.10) in order to establish whether the estimates from reverse survival and relational gompertz models are consistent. Table 4.10 shows that despite the two models having different assumptions, the estimates derived are very close. We can say that Kenyan fertility lies between a low of 4.34 and a high of 4.87. This is inline with other estimates that were done in a period
close to 2009 when the census was done which reported Kenyan fertility to be 4.6 children per woman (KNBS and ICF macro 2010), hence the estimates are consistent.

The study revealed that fertility estimates from the reverse survival model have an added advantage since they can reveal a trend in fertility. As shown in figure 4.6, fertility started stalling in Kenya in the year 2000 to 2007 (TFR 4.9), this supports already existing literature in the country of fertility stall during 1998-2003 with a TFR 4.8 (Westoff, Anne, and Charles 2006; Mutuku, 2013). They explained that the decline of fertility had stalled because of the plateau in contraceptive prevalence and, perhaps more fundamentally, a shift toward wanting more children.

This study revealed that fertility in Kenya is still high with women in rural areas having more children compared to those living in Urban areas. It is important to note that, the greatest contributor of the high fertility in Kenya was education. A possible explanation for this finding is that a little formal education, i.e. primary school level only, does little to change the fertility behaviour of females. The finding that the level of education of females is inversely related to fertility and that the impact of education begins to be felt after basic education underlines the need for at least secondary education for females for any influence on fertility behaviour, and consequently reproductive health to be felt (GSS, 2013). This finding is also supported by other researchers who have studied fertility in Kenya recently. For instance, while studying fertility transition and its determinants in Kenya, Mutuku (2013) found out that secondary and above level of educational attainment is associated with decline in fertility. He found out that women with secondary and above level of education were 23% less likely to transition from third birth to fourth birth compared to women with no education.
CHAPTER FIVE
SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.1 Introduction
This chapter summarizes the research findings, makes conclusions and recommendations. Recommendations are made for policy makers and future researchers. These are based on the research findings.

5.2 Summary
This study set out to estimate fertility levels and trends in Kenya. Specifically, it sought to determine fertility levels and trends in Kenya, determine fertility level by region of residence and level of fertility by place of residence and education level.

To achieve the above objectives, estimation of fertility levels was done using the gompertz and reverse survival models at national level. Fertility at sub-national level was estimated by the relational gompertz model. The study utilized data from the Kenya population and housing census that was conducted during the night of 24/25th of August 2009.

The results at national level revealed that fertility in Kenya has been declining. However, there was a stall in fertility decline during the period 2000-2007. During this period, fertility leveled at 4.9 children per woman. The results obtained by the two models were found to consistent when compared with those from the Brass consistency check. At sub-national level, fertility was high in rural compared to urban. The study also revealed that women with secondary and above level of education had the least fertility compared to those women who had never attended school.

With respect to the first objective of this study, the TFR of Kenya has been reducing though slow. The results revealed a stall in 2000-2007 (TFR of 4.9). Fertility level varied from 4.3, 4.4 and 4.8. Although fertility in Kenya has been reducing, it is still high and it is not reducing at a rate that was earlier expected as this can be seen from the estimates of the reverse survival which clearly show a stall in fertility decline during the period 2000-2007. The drivers of this high fertility are mainly residence and education level. Women from rural areas reported higher fertility than those from urban areas. Analyses reveal that education is inversely proportional to
fertility i.e. the higher the level of education a woman has the less fertility she is likely to have at the end of her reproductive period if the current age specific fertility remain constant.

The inverse relationship observed between educational attainment and fertility suggests the catalytic role of education in fertility decline. Higher levels of formal education lead to delays in marriage and childbearing empower women and afford them the opportunity to take decisions that affect their lives. Educating females up to at least secondary school level will open the door to a whole view of opportunities including a favorable attitude towards small family sizes with its attendant benefits (GSS, 2013). These two factors, education and residence may not be enough to explain why fertility is still high in Kenya, therefore, there is need to examine other general social or economic changes that have recently occurred in Kenya beyond the individual characteristics. For example, the role of the government and international donor support for family planning may very well have contributed to the stall in contraceptive prevalence. The increase in the proportion of women who want more children is more confusing (Westoff, Anne and Charles 2006).

5.3 Conclusion
With regard to method of analysis, this study concludes that, In spite of its simplicity, the reverse survival method of fertility estimation has rarely been used probably because it tends to underestimate fertility levels compared to other methods commonly used in Kenya; the Relational gompertz model and Parity progression ratios. Moreover, analyses revealed that the estimates from reverse survival are equally accurate and consistent besides producing a trend of fertility levels which the relational Gompertz model does not. The method can be applied to a bulky body of existing and easily available population data both modern and past that so far has remained largely under-exploited, and add to the study of fertility levels and trends.
5.4 Recommendations

This section presents recommendations for further research and for policy.

5.4.1 Recommendations for policy

In view of study findings, the study recommends that, the Government in conjunction with the private sector, should put more effort in increasing the proportion of women with secondary and above education. This is because women with secondary and above level of education, had fewer TFR compared to those with none and education plays an important role in changing attitude and behaviour towards reproduction. In addition, higher education will open up better job opportunities thereby making women to have alternative investments besides children. This will ultimately reduce fertility in Kenya.

5.4.2 Recommendations for further research.

This study experienced a number of limitations. One is the limitation on the use of reverse survival model to estimate fertility at sub-national level or at sub groups so that the results can be compared to those from relational gompertz model. Future research should employ the use of the extension of reverse survival method; “the own children method.” This will make comparisons of the estimates at sub-national level possible.

The other limitation pertains to the use of the standard life table in fitting relational model life tables. This study recommends that in order to increase the consistency of estimates from the reverse survival estimates, one should consider developing life tables for Kenya since they will reflect better levels of mortality compared to the standard life table used in this study (Princeton west).
REFERENCES

Centre for Population Studies, London School of Hygiene and Tropical Medicine.


Zaba, B. 1981. “Use of the relational Gompertz model in analyzing fertility data collected in retrospective surveys.” London: Centre for Population Studies, London School of