# Projecting fertility by educational attainment: Proof of concept of a new approach.

By

Presley Ncube



A dissertation submitted to the Faculty of Commerce in partial fulfilment of the requirements of a Master of Philosophy in Demography at the

Centre for Actuarial Research

University of Cape Town

May 2018

## ABSTRACT

The United Nations Population Division publishes fertility projections for all countries in the World Population Prospects (WPP). These are the most widely used projections for planning and policy implementation. Despite a substantial body of literature that suggests education has a significant impact on fertility, these projections do not incorporate changes in the composition of the population by level of education. We therefore propose and implement a method that incorporates education composition change in projecting fertility. We investigate fertility differentials by level of education, then evaluate how education influences fertility independently; and finally, a model is fitted to project fertility rates by education levels. In both cases, the fertility rates by education level are then weighted by the IIASA educational attainment distributions to get the national fertility rates. These national fertility rates are in turn validated against the WPP fertility rates to evaluate how good the proposed method works. Fertility is high among the less educated relative to educated women. Education proves to be an important driver of fertility decline in Southern Africa. The proposed model is a good fit for countries with sufficient DHS data. However, there are other sources of data that are available, for example, the census data but we could not rely on them since they only give summary information. Validation was done to evaluate how good the model is working. This exercise produced consistent results with the observed fertility estimates. The percentage difference between the projected and WPP fertility estimates varied from 1 to 5 percent in Lesotho, Namibia and Zimbabwe. In conclusion, the model can also be used for other countries. Furthermore, education composition change should be considered when projecting fertility since it has proven to be a significant driver of fertility change. Data quality and availability issues were a major limitation to our study and in future should be improved.

# TABLE OF CONTENTS

## LIST OF ABBREVIATIONS

| | |
|---|---|
| AIDS: | Acquired Immune Deficiency Syndrome |
| ASFR: | Age-Specific Fertility Rate |
| BHM: | Bayesian Hierarchical Model |
| DHS: | Demographic and Health Survey |
| HIV: | Human Immunodeficiency Virus |
| IIASA: | International Institute for Applied Systems Analysis |
| ISCED: | International Standard Classification of Education |
| LDHS: | Lesotho Demographic and Health Survey |
| LFS: | Labour Force Survey |
| MAC: | Mean Age at Childbearing |
| MICS: | Multiple Indicator Cluster Survey |
| NDHS: | Namibia Demographic and Health Survey |
| NSO: | National Statistics Office |
| SADHS: | South Africa Demographic and Health Survey |
| SDG: | Sustainable Development Goal |
| SSP: | Shared Socioeconomic Pathway |
| SZDHS: | Swaziland Demographic and Health Survey |
| TFR: | Total Fertility Rate |
| UIS: | UNESCO Institute for Statistics |
| UN: | United Nations |
| UNESCO: | United Nations Education, Scientific and Cultural Organisation |
| UNDESA: | United Nations Department for Economic and Social Affairs |
| WIC: | Wittgenstein Centre |
| WPP: | World Population Prospects |
| ZDHS: | Zimbabwe Demographic and Health Survey |

# LIST OF TABLES

# LIST OF FIGURES

# ACKNOWLEDGEMENTS

Special gratitude goes to Professor Tom Moultrie, my supervisor who made this study a success. Many thanks to him for all the guidance and encouragement to work harder, and all the determination he put towards this work. I also want to thank the two organisations for the scholarships they awarded me to finance my studies and living expenses while studying at the University of Cape Town: The Hewlett Foundation and the National Research Foundation: Freestanding, Innovation and Scarce Skills Development. Special thanks to my family for all the support they gave me all the time.

# 1 INTRODUCTION

To date, fertility projections do not incorporate the effects of a change in the level of education of the population on fertility despite much research that indicates education has a significant effect on fertility change. The fertility projections most widely used are published biennially by the United Nations Population Division in the World Population Prospects (WPP[1]), of which the 2015 report being the latest publication (United Nations Department of Economic and Social Affairs Population Division, 2015). The International Institute for Applied Systems Analysis (IIASA) project the age-education proportions by sex is advance relying on fertility, mortality and migration assumptions from the WPP (Lutz and KC, 2013).

These fertility projections do not incorporate a change in the level of education of the population, which implicitly assumes that the impact of the change is reflected in average change in fertility. This study will therefore investigate if the incorporation of a change in the level of education of the population leads to better project fertility. This application of the educational component in projecting fertility will help to investigate whether allowing for it improves the accuracy of projections. Past and projected educational attainment distributions for all countries are published by the International Institute for Applied Systems Analysis (IIASA[2]) in different scenarios explaining different education policies that can be implemented by individual countries (K.C. and Abel, 2015). These are adapted to investigate the effects of change in the level of education of the population on fertility and validating fertility projections by education level. This study is conducted in five countries in Southern Africa: Lesotho, Namibia, South Africa, Swaziland and Zimbabwe.

## 1.1 Background to the study

This dissertation explores an alternative way of projecting fertility, namely projecting it by educational attainment, in particular countries in Southern Africa. Many authors (Cochrane, Pandey, Pandey *et al.*, 1979, Jejeebhoy, 1995, Cleland and Jejeebhoy, 1996, Bongaarts, 2010, Shapiro, 2017) have argued that level of education plays an important role in fertility decline. However, fertility projections by the UN WPP and several others

[1] United Nations Department of Economic and Social Affairs Population Division. 2015. *World Population Prospects: The 2015 Revision, custom data acquired via website* [Online]. Available: http://esa.un.org/unpd/wpp/DataQuery/ [Accessed 21 February 2016.

[2] Wittgenstein Centre for Demography and Global Human Capital. 2016. *Scenario Definitions* [Online]. Available: http://www.oeaw.ac.at/vid/dataexplorer/ [Accessed 6 June 2016].

(Alkema, Raftery, Gerland *et al.*, 2011, Raftery, Alkema and Gerland, 2014, Miller, 2006) do not consider education composition effects when projecting fertility, thereby, implicitly assuming the impact of change in the level of education of the population on fertility is reflected in the average change in fertility.

Education attainment for countries of the world have been projected by the International Institute for Applied Systems Analysis (IIASA) and are publicly available on their website[3]. Various scenarios are published reflecting different education policies that may be implemented by individual countries. These may be used in studies to investigate the past and future implications of education composition change on fertility dynamics. In addition, these may show us how different education policies lead to different fertility trends.

Therefore, we seek to project fertility by education categories, reconstructing and projecting fertility for these different population subgroups. These estimates are evaluated for plausibility by validating them with the latest WPP fertility estimates. Further sample validation performance may show how well the model projects fertility. This process will utilise a subset of the available data used for the initial parameter estimation and model selection to project fertility and compare the results to the corresponding observed fertility estimates.

### 1.2 Problem statement

Most countries rely on fertility projections in planning and policy development. These projections inform countries about the future demand and supply for food, water and services. Often countries depend on projections published by the United Nations World Population Prospects. The most trusted method of population projections is the cohort-component method, which relies on future fertility estimates. The accuracy of these fertility estimates is thus of some importance. However, the fertility estimates are derived without specifically incorporating educational composition change on fertility dynamics. The question that arises is whether it is possible to reproduce estimates of TFRs similar to those of the WPP and the later DHSs by using a different technique that utilises previous education-status specific TFRs to IIASA projections of proportions of women by education status? Therefore, we incorporate women's

[3] Wittgenstein Centre for Demography and Global Human Capital. 2015. *Wittgenstein Centre Data Explorer Version 1.2*. [Online]. Available: http://www.wittgensteincentre.org/dataexplorer [Accessed 16 February 2016].

educational attainment in projecting fertility under different policies of education, and determine how these changes may affect future patterns of fertility.

## 1.3 Rationale of the study

In this study, we focus on fertility but not the other two demographic components (mortality and migration) because "in most settings and in the long-run fertility is the single most important determinant of population dynamics and growth" (Moultrie, 2013, pg 25). Moreover, data on both migration and mortality tend not to be reliable in developing countries, due to incompleteness of civil registration systems, and at a national level the net contribution of migration to population growth is relatively small compared to fertility and mortality (Timæus, Dorrington and Hill, 2013, Hill and Dorrington, 2013). The study is, however, limited by time and availability of data in these selected countries, therefore we focus on fertility projections.

IIASA produces biennial projections of educational attainment distributions for different education policy scenarios. The study sets out to utilise these data to investigate how different education policies might affect fertility change. It may be useful to consider how fertility may be expected to change when changes in educational attainment over time are considered.

## 1.4 Objectives of the study

The aim of the study is to propose and explore an alternative way of projecting fertility using data on fertility differentials by education collected in surveys in five countries in Southern Africa to contribute to the improvement of the accuracy of fertility projections. However, this exercise demands certain objectives be met. These are presented below.

Initially, we must describe the data required, where they are accessed and quality. Second, what are the characteristics of the population under study. Thirdly, what can be done to ensure that the data are comparable across countries, and presented in the same education categories across data sources.

In addition, we describe how to calculate age-specific fertility rates from the data used. The follow-up question is how does education drive fertility under different education policies and can education policies have an effect independently. Can we then fit a model to project fertility by education level and what parameters will it rely on and what assumptions must be made? Furthermore, are the fitted estimates different from the UN WPP TFRs Once the model is fitted to the countries under study will it prove useful when validated to another country, in this case Kenya. Finally, fertility is projected based on a sample of the data (that is, the early Kenyan DHS) to check how well the measure captures further changes (to predict fertility in the latest DHSs).

### 1.5 Research outline

This dissertation is presented in five chapters. Chapter Two presents a review of the literature on fertility studies, UN WPP fertility projections and how education drives fertility change. Chapter Three describes the three data sources (DHS, IIASA and WPP) used, presents the background characteristics and evaluation of DHS data, and presents the methodology adopted to project fertility by educational attainment. The results from the implementation of the methods are presented and described Chapter Four. Conclusions to the research are presented in the final chapter, giving a discussion of results, summary of the findings and recommendations for further studies.

# 2 LITERATURE REVIEW

This chapter outlines the relevant literature to our study. We begin by describing the determinants of fertility decline in general and outline the role of education on fertility change. This provides the essential justification for the approach of projecting fertility decline by educational attainment. We then review the United Nations (UN) fertility projections which are published on the World Population Projections (WPP) report every two years.

## 2.1 Determinants of fertility decline

Fertility in the sub-Saharan Africa region was the highest in the world in the 1970s (Jolly and Gribble, 1993, Shapiro and Gebreselassie, 2013). While fertility has fallen across the region since then (Caldwell, Orubuloye and Caldwell, 1992, Blacker, 1994, Alkema, Raftery, Gerland *et al.*, 2011), it remains the region with the highest overall fertility (United Nations Department of Economic and Social Affairs: Population Division, 2015), where at present (2015-2020) the medium variant for TFR is 4.75 children per woman (Shapiro, 2017). We therefore need to explain the factors that have driven fertility decline over the years to have a clear understanding of fertility and the factors that have brought about this change.

The study of fertility determinants goes back to the work of Davis and Blake (1956), in which they argued that fertility change was influenced by two types of determinants: the distal and the proximate determinants. The former is characterised by socioeconomic, cultural and environmental factors that affect fertility change through proximate determinants. The examples of distal determinants include educational attainment, participation of women in the labour force, religion, cultural and social structures, etc. The proximate determinants are those that immediately affect the reproductive process of women in biological terms. The process of reproduction involves three steps: (1) sexual intercourse, (2) conception, (3) gestation and parturition. Therefore, the discussions of proximate determinants will be clustered by how they affect the reproductive process.

These determinants were classified into eleven factors, later classified into eight by Bongaarts (1978) and finally into four proximate variables by Bongaarts and Potter (1985). Table 2-1shows the proximate determinants which were trimmed from eleven to four. Bongaarts and Potter (1985) argue that the other seven were relatively invariant across human populations and were difficult to measure.

Likewise, the eleven determinants identified by Davis and Blake (1956) fall into three major classes according to how they affect reproduction: the first six determinants affect the exposure to sexual intercourse, while the determinants seven to nine affects conception and finally the factors ten and eleven affect gestation and successful birth deliverance.

**Table 2-1 Direct determinants of fertility change**

| *Davis and Blake (1956)* | *Bongaarts (1978)* | *Bongaarts and Potter (1985)* |
|---|---|---|
| 1. Age of entry into sexual unions | Proportion of women married or in sexual unions | Contraception |
| 2. Proportion of women never entering sexual unions | Contraception | Induced Abortion |
| 3. Amount of reproductive period spent after or between unions | Induced abortion | Lactation amenorrhea |
| 4. Voluntary abstinence | Lactation amenorrhea | Fecundability |
| 5. Involuntary abstinence | Frequency of intercourse | |
| 6. Coital frequency | Sterility | |
| 7. Fecundity or infecundability, as affected by voluntary causes | Post-partum abstinence | |
| 8. Use or non-use of contraception | Duration of the fertile period | |
| 9. Fecundity or in-fecundity as affected by involuntary causes | | |
| 10. Foetal mortality from involuntary causes | | |
| 11. Foetal mortality from voluntary causes | | |

In Table 2-1, the first three determinants affect the formation and termination of sexual unions, while the fourth and the sixth accounts for the exposure to intercourse. Late entry into sexual unions by women has been shown to have a negative impact on fertility. On the other hand, voluntary and involuntary abstinence and the frequency of sexual intercourse directly affect the way in which people in sexual unions indulge in conjugal activities. These determinants are very complex to quantify and therefore data on them tends to be unreliable (Bongaarts, 1978).

The seventh to ninth determinants affect fertility through prevention or avoidance of conception voluntarily and involuntarily, either through use of contraception (or sterilisation, abstinence, etc.) or from primary infertility. The last two determinants identified by Davis and Blake (1956) affect fertility in that the pregnancy of a woman is voluntarily or involuntarily terminated.

One of the most important advantages of studying the effects of proximate determinants on fertility change is that it gives an understanding of how socioeconomic and environmental factors operate to influence fertility change. The effect of these determinants on fertility may either be positive or negative, and the actual fertility rates are dependent on the net balance of the effects between different proximate determinants. If, for example, women effectively utilise contraceptive measures, then contraception will have a negative effect on fertility change. However in the absence of contraceptive use fertility tends to be high in developing countries (Asiimwe, Ndugga and Mushomi, 2013). Researchers must be careful in interpreting the absence of contraceptives as there is lack of substantial experimental evidence that their absence contributed to such high levels of fertility in developing countries but in general, lower use of contraception implies higher fertility. However, this may be due to the influence exerted by other proximate determinants. Hindin (2000) argue that since the introduction of family planning in Zimbabwe in 1985, fertility levels declined consistently from 6 in 1985 to 4 children per woman by the 1994 Demographic and Health Survey (DHS). Much of Zimbabwe's success in reducing the fertility rates has been attributed to the role played by the government and non-governmental organisations in educating women about the use and offering the most effective contraceptives.

Temporary infertility after birth results from lactational amenorrhea, women who tend to breastfeed longer, reduce the risk of conception. A study conducted in West African countries (Jejeebhoy, 1995) ascertained that women who breastfeed for longer periods have fewer children. Breastfeeding for longer periods widens the child spacing gap (Lesthaeghe, Ohadike, Kocher and Page, 1981), which in turn reduces the number of children a woman can bear during her entire reproductive lifespan (15-49 years).

The distal determinants must affect fertility through at least one of proximate determinants. For example, women's educational attainment may be positively correlated to fertility change if educated women tend to breastfeed their children for shorter periods alone (Cochrane, Pandey, Pandey *et al.*, 1979), thus increasing the chances of getting pregnant soon, since their menstrual cycle will have restored soon after they stop breastfeeding, assuming they do not use other birth control measures. It may also be negatively correlated with fertility if women who are more educated tend to have better access to and use the most effective methods of contraception (pills, IUD and condoms). In addition, more educated women are better informed about these measures and are more likely to use the modern health care for medical problems such as sexually transmitted diseases, which may lead to fecundability (Jolly and Gribble, 1993). The net effect of education on fertility change can be positive, negative or insignificant, depending on how the positive and negative effects contribute to alter the proximate determinants (Bongaarts, Frank and Lesthaeghe, 1984).

Besides education, there are other distal determinants that affect fertility change such as cultural and social structures. Caldwell and Caldwell (1987) ascertain that cultural practices such as post-partum abstinence and attitudes to marriage affect fertility change. Most African cultures urge women to refrain from having sex while they are still breastfeeding for a period up to three years and a "common resort to terminal female abstinence as women become grandmothers" (Akmam, 2001, pg 77). This is done to increase the chances of child survival, thereby increasing child spacing, which in turn may mean that women tend to have fewer children within their entire reproductive lifespan (15-49 years) compared to the case where this abstinence is not practised. The abstinence tempers one variable in the reproductive process of women, that is the exposure to sexual intercourse. However, in polygamous societies men tend to have access to other wives, thereby maintaining fertility rates at high levels (Bongaarts, 1978).

## 2.2 Links of Education to fertility

Among the myriad distal socio-economic determinants of fertility, education has received the most attention from researchers (Jejeebhoy, 1995, Martin, 1995, Lutz and Goujon, 2003, Akmam, 2002, Chang, 1996, Cochrane, Khan and Osheba, 1990, Easterlin, 1983, Cochrane, Pandey, Pandey *et al.*, 1979, Bongaarts, 2010, Shapiro, 2017). However, as a distal determinant of fertility education must act through the proximate determinants as described earlier. In this section we set out mechanisms which link education and fertility outcomes.

Easterlin (1983) argue that education affects fertility change through two channels, that is (1) demand and supply of children, and (2) the cost of fertility regulation. Education influences the supply of children indirectly and operates by interacting with the proximate determinants, such as breastfeeding, postpartum abstinence, child mortality and age at marriage. On the other hand, it has an impact on the demand for children indirectly through the "desired family size, son preference, labour contributions of offspring during childhood, children as sources of prestige, economic and time opportunity costs of raising children and children as old age support" (Jejeebhoy, 1995, pg 19).

Shapiro (2017) investigates the pathways in which education is linked to fertility using DHS data in 30 sub-Saharan countries with years of schooling as an indicator for educational attainment in women. He controlled for the type of place of residence and religion in his regression models. The findings were that better-educated women had preferences for smaller number of children and were less likely to be in sexual unions contributing to lower supply of children. Highly educated women were more likely to use modern contraceptives relative to their lesser-educated counterparts. The increased educational attainment over time and reduced child and infant mortality were shown to have contributed substantially to fertility decline in all countries.

Furthermore, extensive investigations reveal that in general educated women tend to marry later than their uneducated counterparts (Cochrane, Pandey, Pandey *et al.*, 1979, Bongaarts, 2010), and consequently tend to have their first birth when they are older (Jejeebhoy, 1995). This decreases their exposure to childbearing. Cleland and Jejeebhoy (1996) in a study conducted in South Asian countries found that women with higher education got married two to five years later as opposed to uneducated women, and fertility was lower among highly educated women. In societies where fertility is constrained almost exclusively within marriage, later marriage has a suppressive effect on fertility.

More educated parents often feel the need for their children to attain higher education levels than they had or have higher standards of health care, feeding and housing (Easterlin, 1983). Thus educated women perceive these costs to be higher and there has to be trade-off between costs of raising more children with a possibility of compromising the need for them to be well taken care of, and raising a few who can be well taken care of, which assures a better life for them (Cleland and Jejeebhoy, 1996). In the sub-Saharan region, fertility and demand for children has fallen at the same time as education increased. As both Cleland (2002) and Bongaarts (2010) have argued, educational background is strongly negatively correlated with desired family size.

Large numbers of educated women are more likely to be active in the labour force and therefore spend less time at home (Al Riyami, Afifi and Mabry, 2004). The majority of educated women would rather have fewer children than having to leave the labour force to raise their children (Chaudhury, 1978). This opportunity cost tends to force women to limit the number of children they have.

Education often influences the way women behave; the major influential behaviour connecting education to fertility is through the use of contraceptives (Chaudhury, 1978, Cochrane, Pandey, Pandey *et al.*, 1979). The findings by Cleland and Jejeebhoy (1996) note that the use of contraception rises monotonically with increasing education, and in most sub-Saharan African countries the link takes the form of increased preference to the most efficient modern contraceptive methods. Hindin (2000) argues that education in Zimbabwe has had a significant impact on contraceptive use. His findings show that women who have completed primary education were significantly more likely to use contraceptives compared to uneducated women, while those with at least secondary education were twice as likely to use contraceptives as opposed uneducated women. In this case, education is a distal determinant of fertility, which affects fertility change through the contraceptive use (proximate determinant).

Swartz (2009) argues that poverty and a lower level of education among African women persisted in South Africa in the post-apartheid era, while fertility has continued to decline slowly. This is in contrast with the rest of the sub-Saharan countries, as high poverty usually goes hand in hand with high fertility levels. He goes on further to outline that populations trapped in high poverty and inequality are often less educated. Swartz found that even though most women were trapped in poverty and characterised by lower education relative to their fellow counterparts in sub-Saharan Africa, fertility was lower due to the use of contraceptives as a survival strategy.

### 2.2.1 IIASA education scenarios

To date, quantitative studies on fertility trends do not explicitly factor in the effects of educational composition change. Education composition change is the increase or decrease in the proportion of women in specified education levels over time. The International Institute for Applied Systems Analysis (IIASA) has recently produced population projections by educational attainment (Lutz and KC, 2013). According to the authors, the exercise was undertaken for three reasons: (1) to update the previous projections with the more recent data from surveys and censuses; (2) to increase the number of countries they provide estimates for in order to have a clear vision worldwide on levels of educational attainment and their potential future influence; (3) to increase the education attainment categories from four to six in order to allow for greater precision in incorporating levels of education across the globe.

The sources of data used by IIASA to generate educational distributions for the countries studied in this dissertation are summarised in Table 2-2. In all countries in Southern Africa, with exception of South Africa, IIASA relied on data from Demographic and Health Surveys (DHS). In South Africa, in the absence of a reliable current DHS, data from the 2007 Community Survey was used. However, not all surveys represent the sample across sex, age and territory. And While the data from surveys are the most common source of data on fertility and fertility outcomes, these data are often far from perfect in their coverage or their context. Schoumaker (2014) notes that the poorest and least educated households are often located in remote areas and hence households may be omitted with greater frequency, running the risk of compromising the representativeness of the sample to the whole population. To address coverage errors, post-sampling weighting is used; while extensive data and cross-validation at the collection stage was used to reduce the content errors.

**Table 2-2 IIASA data sources on educational attainment in Southern Africa**

| *Country* | *Data Type* | *Source* |
|---|---|---|
| Lesotho | DHS 2009 | MEASURE DHS |
| Namibia | DHS 2006-07 | MEASURE DHS |
| South Africa | Community survey (2007) | NSO |
| Swaziland | DHS 2006 | MEASURE DHS |
| Zimbabwe | DHS 2005-06 | MEASURE DHS |

Source: Speringer, Goujon, KC *et al.* (2015)

The data on educational attainment was categorised into six education categories as shown in Table 2-3. Different countries have different educational systems with differing numbers of years of study at primary and secondary levels, and without an attempt at harmonisation, data from different countries are frequently not comparable. IIASA used the UNESCO International Standard Classification of Education (ISCED) first published in the 1970s, and later revised in 1997 and again in 2011 to derive harmonised measures of population distribution by levels of education. However, country-specific ISCEDs using the 2011[4] classification system are not yet available for all the countries. Therefore, IIASA used the 1997 ISCEDs.

**Table 2-3 Educational attainment categories**

| *IIASA 2015* | *ISCED[5] 1997* | *IIASA categories (before 1997)* |
|---|---|---|
| No education | No level ISCED 0<br>Incomplete grade 1 of ISCED 1 | No education |
| Incomplete primary | Incomplete ISCED 1 | Primary |
| Completed primary | Completed ISCED 1<br>Incomplete ISCED 2 | |
| Completed lower secondary | Completed ISCED 2<br>Incomplete ISCED 3 | Secondary |
| Completed upper secondary | Completed ISCED 3<br>Incomplete ISCED 4 or 5B | |
| Post-secondary | Completed ISCED 4 and 5B<br>ISCED 5A and 6 | Tertiary |

Source: Bauer, Potancokova, Goujon and KC (2012)

IIASA publishes seven assumptions for educational attainment distributions varying fertility, mortality and migration levels against different education policies using a Delphi-technique (KC and Lutz, 2014). These are commonly known as the Shared Socio-economic Pathways (SSP). The SSP scenarios have been tailored to capture alternative future developments with respect to socioeconomic encounters to both climate change mitigation and adaptation. Of these scenarios only two are reasonably applicable in Southern Africa (KC and Lutz, 2014), given the short projection interval considered in the present study: the SSP2-GET[6] and the SSP2-CER. The SPP2-GET scenario is the most likely for all countries in the region as it combines experiences of countries characterised by medium fertility, medium migration and medium mortality and that it assumes a Global Education Trend (GET). The SSP2-CER is the same as

[4] The only difference between the ISCED 1997 and 2011 is the extension of the categories covering tertiary education from 2 to 4 levels. It must, however, be noted that since the "post-secondary" encompasses all these tertiary education groups, the ISCED 2011 remains in line with the ISCED 1997.

[5] ISCED 0-preceding primary education, ISCED 1-primary education in accordance with basic education, ISCED 2-lower secondary or second stage of basic education, ISCED 3-upper secondary education, ISCED 4-post-secondary non-tertiary education, ISCED 5-first stage of tertiary education (either 5A-longer research specific, 5B-short technical occupation), ISCED 6-second stage of tertiary education leading to advance research (PhD or Doctorate).

[6] The detailed explanation of all the scenarios is given by the Wittgenstein Centre for Demography and Global Human Capital. 2016. *Scenario Definitions* [Online]. Available: http://www.oeaw.ac.at/vid/dataexplorer/ [Accessed 6 June 2016].

that for SSP2-GET except it assumes a Constant Enrolment Ratio (CER) education trend. The GET scenario is derived by assuming that the developing countries will follow the average path of education (or school) expansion that other developed or further advanced countries have already experienced in the past, while the CER scenario assumes that for each country the recently observed rates of school enrolment will remain constant or frozen at the current levels as we progress into the future.

The population is projected by age, sex and education by different SSP scenarios using the multi-state projection model described by Lutz and Goujon (2001), which is basically the cohort-component method applied beyond age and sex and the method requires components of demographic change to be projected in advance. From these projected population numbers, the distributions by education level in each age group can be derived for both men and women. We specifically focus on the distribution of women by education level in each five-year age group (15-19, 20-24, …, 45-49). These distributions were downloaded from their website[7], for each country, which are available from 1970 to 2100.

## 2.3 UN WPP fertility projections

The United Nations (UN) Population Division publishes their population projections in the World Population Prospects (WPP). These are released approximately every two years in order to incorporate new data as they become available and allow reformulation of detailed assumptions about future paths of fertility, mortality and migration (United Nations Department of Economic and Social Affairs: Population Division, 2015). The past estimates in the most recent revision run from 1950 to 2015, and projections run from 2015 to 2100.

The UN WPP projections have evolved over the years from the use of non-probabilistic approaches to include the current probabilistic approach: The Bayesian Hierarchical Model (BHM). Three scenarios are normally explored: (1) a benchmark scenario, which is regarded as the best guess about the future trend in fertility; together with (2) low and (3) high scenarios bracketing the benchmark projection. These are usually used to provide some measure of uncertainty around the benchmark scenario. Yet this approach has been criticised on multiple grounds by many demographers and practitioners (Lutz and Goujon, 2003, Alkema, Raftery, Gerland *et al.*, 2011, Miller,

---

Wittgenstein Centre for Demography and Global Human Capital. 2015. *Wittgenstein Centre Data Explorer Version 1.2.* [Online]. Available: http://www.wittgensteincentre.org/dataexplorer [Accessed 16 February 2016].[7]

2006, Lee and Tuljapurkar, 1994) leading to the development of recent methods described below.

The Random Country Method (RCM) was developed by Miller (2006) and based on the collective experience of 192 United Nations (UN) member countries. The TFR estimates were produced for five-year periods using the procedure described below. The RCM method starts by recording the initial or the base year TFR for a specified country: for example, if the base year TFR was 6 children per woman in Zimbabwe in 1980, the next step would be to select a sample of 100 countries with the TFR which is closest to 6 children per woman at any point in time. From this sample of 100 countries, one country's TFR is selected at random, and its fertility over the next five-year period is used to set the level of fertility in the target country. Thus, the method uses a sample of convenience to predict future TFR for five-year periods. This process is repeated until the required projection length is attained. For example, if the projection is for 20 years from 1980, this procedure will be performed to produce the results until 2000 for the five-year intervals.

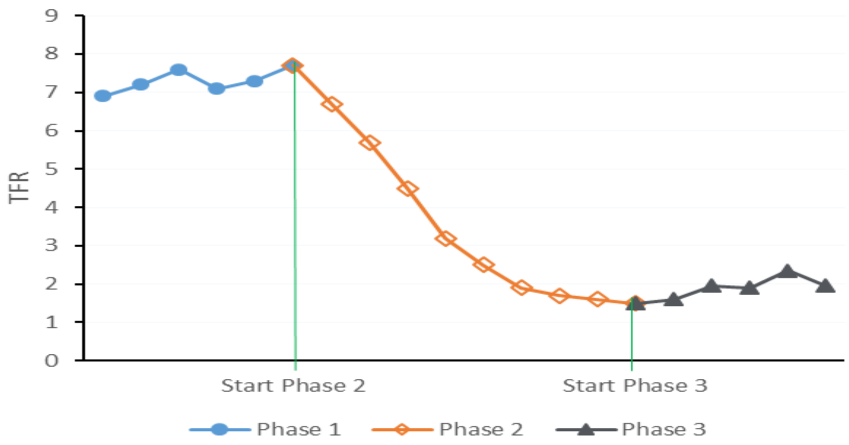
The assumption underlying this approach is that the current level of fertility is only influenced by the most recent period of fertility and furthermore assumes that countries experience similar economic, social and cultural conditions and forces. It thus ignores the uniqueness of countries, and for this reason we can find huge differences in TFR estimates between this approach and others.

In addition, estimates are often biased in forecasting fertility (Bauer, Potancokova, Goujon and KC, 2012): when fertility is high in the current period they tend to result in the projection of a continued high level of fertility in subsequent periods, and vice versa. Second, they provide an inconsistent measure of uncertainty around the fertility estimates (Miller, 2006), which are characterised by a narrow bound in the early years of projections and bounds that are too wide in the later years of the projection period. Faced with such criticisms, an alternative approach was developed by Alkema, Raftery, Gerland *et al.* (2011) and is currently used by the United Nations and described in Section 2.3.1.

### 2.3.1 Bayesian Fertility Projection Model

The current approach used in the WPP to project fertility follows the approach set out by Raftery, Alkema and Gerland (2014). The projection methodology for this model is based on demographic transition theory, which states that societies will move from a pre-transition phase (Phase 1) characterised by high fertility and mortality to a posttransition phase (Phase 3) where both are low (Dyson, 2011). There is general agreement (Kirk, 1996, Teitelbaum, 1975, Chesnais, 1992, Cai, 2010, Rogers and Hackenberg, 1987, Dyson, 2011) that fertility change occurs in three distinct phases, as illustrated in Figure 2-1. This fertility projection model is only applied to phases two (2) and three (3), by the UN Population Division, since all the member states of the United Nations have passed this pre-transition phase and therefore would make no sense to model it (United Nations Department of Economic and Social Affairs: Population Division, 2015).

**Figure 2-1 Three major phases of the model for total fertility**




Source: Raftery, Alkema and Gerland (2014) page 61

Countries were grouped by the United Nations Population Division (2015) into three groups: high fertility-classified as phase one (1) countries although all member states have passed this stage; medium fertility-those in phase two (2) who are undergoing a systematic fertility decline; and low fertility countries-in phase three (3) where the total fertility is fluctuating about replacement-level fertility and below. Phase two (2) countries were modelled by a double logistic decline function, which is a function of its level as shown by equation (1). The parameters ($\delta^c$) of this function were estimated using a Bayesian Hierarchical Model described in detail by Raftery, Alkema and Gerland (2014) which resulted in parameters suited for each country.

$$f_{c,t+1} = f_{c,t} - r(f_{c,t} / \delta^c) + \varepsilon_{c,t} \tag{1}$$

The symbols are defined below:

$f_{c,t}$ -is the total fertility rate for country '*c*' at time '*t*' in five-year periods,

$r(f_{c,t} / \delta^c)$ -five-year decrement of the TFR,

$\delta^c$ -independent and identically distributed normally country-specific parameters,

$\varepsilon_{c,t}$ -error term which is normally distributed.

The country-specific fertility distributions were informed by historical trends within each country, and the variations in distributions across other countries were used to determine the level of uncertainty around the projected fertility estimates, based on what other countries experienced when they were at similar levels of fertility. The model is termed hierarchical because, in addition to the country-specific information that we have, a second level of information (the worldwide experience according to the information of all countries) is used to inform the statistical distributions for the country-specific parameters of the double-logistic function. The model produced reasonable results when reconstructed and compared to observed fertility (Alkema, Raftery, Gerland *et al.*, 2011) even in cases where countries had limited information about their past fertility trends. However, it can be argued that if there is limited data, the extent to which these estimates are reasonable is subject to debate.

The second aspect of the model involves the projection of fertility for countries that were already in phase three (3) of the demographic transition. These countries are characterised by low fertility levels, and the TFR decreased to around replacement-level fertility. In the long run, the TFR is assumed to converge towards and fluctuate around replacement-level, as proposed by Lee and Tuljapurkar (1994) with the motivation from observed increments of TFR from below replacement-level fertility. The process is modelled with a first-order autoregressive time-series (AR (1)) model for countries in phase three. This model has a mean fixed at the approximate replacement-level fertility (2.1 children per woman). This is summarised in equation (2), where "$\rho$" is the autoregressive parameter: $|\rho|<1$.

$$f_{c,t+1} = f_{c,t} + (1-\rho)(\mu - f_{c,t}) + \varepsilon_{c,t} \quad (2)$$

The TFR is expected to increase if it is below replacement-level fertility, and vice versa. The smaller the value of the autoregressive parameter, the more quickly the TFR will increase towards replacement-level fertility. It is evident that the AR(1) model predicts recovery from below replacement-level fertility, as does the Bongaarts and Feeney (1998) work on tempo and quantum effects of fertility. For this reason, the method is commendable as it captures this phenomenon. Only a few countries have reached this phase in the world, and the majority of them achieved this in recent years (Alkema, Raftery, Gerland *et al.*, 2011). Hence the AR (1) model is the ideal method for now because there is currently insufficient data to fit higher order autoregressive models for higher precision.

## 2.4 Conclusion

Having reviewed the literature, it is apparent that education is an important determinant for fertility. Therefore, we will incorporate IIASA educational attainment distributions to project fertility in several countries. Adopting these distributions to project fertility by education levels may raise new insights if they are validated against the UN WPP fertility estimates while controlling for heterogeneity that arises from different education levels.

# 3 DATA AND METHODS

This section starts by describing the datasets used in the study, then shifts the focus to the background characteristics and evaluation of the Demographic and Health Survey (DHS) data. Finally, we outline the methods used to build a model to predict fertility into the future.

## 3.1 Description of datasets

Descriptions of the three datasets used are given in this section. We start by giving a description of the Demographic and Health Survey (DHS), move on to the International Institute for Applied Systems Analysis (IIASA) population projections by educational attainment and finally the United Nations (UN) World Population Prospects (WPP) data.

### 3.1.1 DHS data

DHS data are preferred in many demographic applications since they are considered as the most reliable of data in developing countries, where civil registration systems are mostly incomplete, and which typically do not collect sufficient background characteristics such as educational attainment. Censuses only collect summary information on fertility and are conducted only once a decade in most developing countries. Therefore, these are less preferred to DHS data, which offer in-depth information on fertility and it is cheaper to finance surveys compared to censuses.

The DHS is designed to collect data on fertility, family planning, reproductive health, nutrition, child health, HIV/AIDS and many other characteristics. For this dissertation, women of reproductive age (15-49 years) are the focus. All DHS utilise a minimum of two questionnaires: the household and women's questionnaire. The household questionnaire is used mainly to identify women, men and children eligible for individual interviews, while the women's questionnaire provides data of women's background characteristics (age distribution, literacy, religion, etc.), reproductive history, fertility preferences and more.

The analysis is performed on fertility data collected as part of the women's questionnaire from the DHS website[8] in STATA format, specifically using the data from women aged 15-49 years about their detailed birth histories from selected Southern

[8] The Demographic and Health Surveys Program. 2016. *Datasets* [Online]. Available: https://dhsprogram.com/data [Accessed 20 January 2016].

African countries (Table 3-1). The sample sizes of DHSs (in general, as well as those selected for analysis) have tended to increase over time. As such, the level of uncertainty around estimates produced is reduced. This is, however, not always the case, for example, a slight reduction in the sample size was observed between the last two DHSs (2006-07 and 2013) in Namibia. The same was observed between the 2009 and 2014 DHSs in Lesotho and between 1994 and1999 DHSs in Zimbabwe. Table 3-1 shows the response rates in the region, which are generally above 90 percent, with the highest response rate seen in Lesotho in 2009 and the lowest in Zimbabwe in the 2005-06 DHS.

The number of eligible women identified in the 2014 Lesotho DHS is much lower than the 2009 LDHS (Ministry of Health [Lesotho] and ICF International, 2016). This explains the drop in the sample size of the 2014 LDHS. The response rates are also lower compared to the 2009 LDHS. Similarly, in Namibia response rates drop in the most recent survey.

**Table 3-1 Southern Africa DHS data**

| *Country* | *Survey year* | *Successful interviews (sample size)* | *Response rate (%)* |
|---|---|---|---|
| Lesotho | 2004 | 7,095 | 94.3 |
| | 2009 | 7,624 | 97.9 |
| | 2014 | 6,621 | 94.0 |
| Namibia | 1992 | 5,421 | 92.7 |
| | 2000 | 6,755 | 92.4 |
| | 2006-07 | 9,804 | 94.7 |
| | 2013 | 9,176 | 92.3 |
| South Africa | 1998 | 11,735 | 95.2 |
| Swaziland | 2006-07 | 4,987 | 94.1 |
| Zimbabwe | 1988 | 4,201 | 94.0 |
| | 1994 | 6,128 | 95.6 |
| | 1999 | 5,907 | 95.2 |
| | 2005-06 | 8,907 | 90.2 |
| | 2010-11 | 9,171 | 93.3 |

Source: Derived from country-specific DHS reports

The potential of non-response bias within each of these surveys is less than 10 percent. This would be an issue if distinct differences exist between people who responded and those who did not respond to the survey. Otherwise, this is generally not an issue: the higher the response rates, the less likely the error from the non-response bias. DHS includes a sampling weight variable to correct for over-or undersampling of women in the sample in each survey because of the differential response rate or sample design (Rustein and Rojas, 2006), and to ensure that the data are representative of the national population as possible. Therefore, in severe cases where the response rates are low but the sample is perfectly distributed the bias will be low, although the validity (generalisability) will be compromised.

The DHS collects full birth history data starting from the first child a woman has given birth to. Three pieces of information are required to compute fertility rates from the DHS data: (1) date of birth of each mother (whether she has ever had given birth or not); (2) date of birth of each child; and (3) the date of the survey. All dates are expressed in Century Month Code (CMC), which is defined as "the number of months since January 1900. It is calculated by multiplying twelve by the difference between the year of the event and 1900. The month of the event is added to the previous result" (Rustein and Rojas, 2006, pg 15). From these data we use the *tfr2* STATA macro programmed by Schoumaker (2013) to calculate the ASFRs counting the events and measuring the exposure in five-year age groups for defined periods.

The standard recode datasets for women also have information about their education. The two questions of many that are used here are: (1) what is the highest education level you attended (primary/secondary/tertiary), and (2) what is the highest grade (standard/form/year) you completed at that level? The information about these responses is found on the variables *v106* and *v107* respectively; and it is from these two variables we generate the six educational categories used later in this research. However, in DHSs carried out prior to 2009 in the region, those who attended less than a year were considered to have attained one year in that level. To deal with that we adjusted those who had attained one year of education to a lower level.

Although the DHS is the most used source of fertility data in the region it also has its weaknesses: data collection is carried out over extended periods of about six months on average, and it takes up to two years before the final report and data are released. In addition, women are expected to recall all births they had in their entire reproductive lifespan, which may have serious consequences of underestimating current fertility if women happen to shift births further back in time to avoid further questions on children born five or less years before the survey. Omission of recent births also has the effect of underestimating current fertility; although DHS is quite good at trying to avoid this, it is seldom problematic in estimates generated.

### 3.1.2 International Institute for Applied Systems Analysis (IIASA) data

The full description of the IIASA data was given in section 2.2, and we use the data to investigate how different education policies lead to different education fertility trends. Two scenarios used, both assuming average demographic characteristics of fertility, migration and mortality. The SSP2-CER (current enrolment rates) scenario assumes current education levels prevail till the end of the projection period. The SSP2-GET (global education trend) assumes that education will change in a manner experienced by developed countries who are at the end of their demographic transition. These assumptions are downloaded from the Wittgenstein Centre for Demography and Global Human Capital (2015) website in the form of age-education proportions for women of reproductive age and described in the section 3.3.

### 3.1.3 World Population Prospects (WPP) data

The focus in this study is on fertility projections that cover the period from 2015 to 2030, which are downloaded from their webpage[9]. As described in the previous chapter, the WPP uses a probabilistic method to model fertility change for countries in the second and third phases of their fertility transition. It assumes that fertility in the long run will decline from high levels until replacement level of 2.1 children per woman and then fluctuate around these levels. However, this may not be the case, as fertility may not decline smoothly as this model suggests and does not capture or allow for situations where fertility can stall for a while, for example, in sub-Saharan Africa, where several studies have highlighted stalls (Johnson, Abderrahim and Rutstein, 2011, Bongaarts, 2006, Westoff and Cross, 2006, Shapiro and Gebreselassie, 2013, Ibisomi, Williams, Collinson and Tollman, 2014).

The reference date for each estimate is the 1st of July, for example, the TFR for 2010 in WPP estimates is applicable from 1 July 2010 to 30 June 2015. These estimates are also available for single year periods and are applicable by calendar year, for example, TFR for 2011 is applicable from 1 January to 31 December 2011. However, for this study we rely on five-year fertility estimates for reasons explained later. The population is projected by age and sex only, thus neglecting other forms of population

[9] United Nations Department of Economic and Social Affairs Population Division. 2015. *World Population Prospects: The 2015 Revision, custom data acquired via website* [Online]. Available: http://esa.un.org/unpd/wpp/DataQuery/ [Accessed 21 February 2016.

heterogeneity, this implicitly assumes that other distinguishing factors like education do not have a direct impact on how populations change.

### 3.2 Background characteristics and evaluation of DHS data

This section describes the characteristics of women interviewed in DHS surveys in each country. Table 3-2 shows the educational composition of women in each country over time. In countries with more than one data set, substantial change in educational attainment is evident. A general rise in the proportion of women attaining post-secondary education is seen in all countries across the region with Swaziland and Namibia having higher proportions of women in this category (about 8 and 7 percent respectively).

**Table 3-2 Percentage educational composition of the interviewed women**

| *Country* | *DHS year* | *No education* | *Incomplete primary* | *Complete primary* | *Lower secondary* | *Upper secondary* | *Post-secondary* |
|---|---|---|---|---|---|---|---|
| Lesotho | 2004 | 7.9 | 27.1 | 46.1 | 11.5 | 6.1 | 1.4 |
| | 2009 | 2.2 | 26.3 | 46.7 | 12.6 | 7.8 | 4.5 |
| | 2014 | 1.3 | 19.2 | 47.0 | 15.0 | 11.0 | 6.4 |
| Namibia | 1992 | 17.8 | 36.8 | 11.0 | 20.0 | 12.8 | 1.6 |
| | 2000 | 12.0 | 20.1 | 12.0 | 41.3 | 11.9 | 2.7 |
| | 2006-07 | 8.0 | 18.6 | 17.9 | 38.2 | 11.6 | 5.7 |
| | 2013 | 7.4 | 16.2 | 7.5 | 46.0 | 16.0 | 6.9 |
| South Africa | 1998 | 7.6 | 16.6 | 9.4 | 44.1 | 21.0 | 1.3 |
| Swaziland | 2006-07 | 9.1 | 20.8 | 32.6 | 16.7 | 13.0 | 7.8 |
| Zimbabwe | 1988 | 13.7 | 34.1 | 26.3 | 11.9 | 13.2 | 0.8 |
| | 1994 | 13.4 | 25.0 | 28.9 | 16.7 | 14.9 | 1.1 |
| | 1999 | 7.5 | 20.6 | 28.4 | 19.2 | 22.0 | 2.2 |
| | 2005-06 | 4.4 | 14.3 | 25.2 | 20.7 | 32.6 | 2.9 |
| | 2010-11 | 2.5 | 11.4 | 24.9 | 21.7 | 35.4 | 4.0 |

Table 3-3 shows the age composition of women of reproductive age in each country; and as expected, the age structure does not change much over the period covered by the DHS.

**Table 3-3 Percentage age distribution of interviewed women**

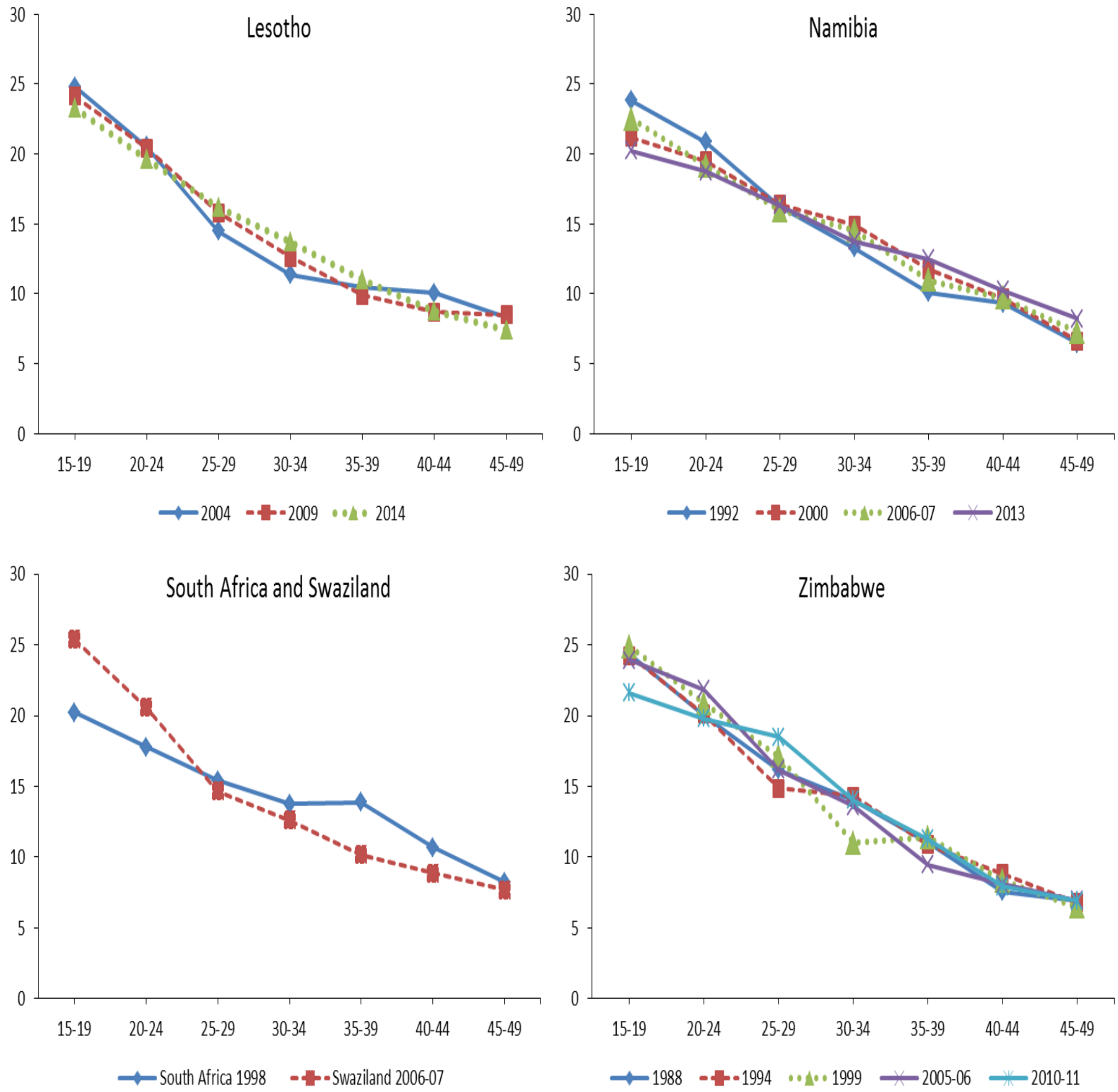
| Country | DHS year | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 |
|---|---|---|---|---|---|---|---|---|
| | 2004 | 24.8 | 20.5 | 14.5 | 11.4 | 10.4 | 10.1 | 8.3 |
| Lesotho | 2009 | 24.1 | 20.4 | 15.8 | 12.6 | 9.9 | 8.7 | 8.5 |
| | 2014 | 23.3 | 19.6 | 16.2 | 13.7 | 11.0 | 8.8 | 7.4 |
| Namibia | 1992 | 23.8 | 20.9 | 16.2 | 13.3 | 10.1 | 9.3 | 6.4 |
| | 2000 | 21.2 | 19.5 | 16.4 | 14.9 | 11.7 | 9.7 | 6.6 |
| | 2006-07 | 22.5 | 19.1 | 15.9 | 14.5 | 11.0 | 9.7 | 7.3 |
| | 2013 | 20.2 | 18.7 | 16.3 | 13.8 | 12.5 | 10.3 | 8.2 |
| South Africa | 1998 | 20.2 | 17.8 | 15.4 | 13.8 | 13.9 | 10.7 | 8.2 |
| Swaziland | 2006-07 | 25.4 | 20.6 | 14.7 | 12.6 | 10.2 | 8.9 | 7.7 |
| Zimbabwe | 1988 | 24.3 | 20.0 | 16.2 | 14.0 | 11.0 | 7.6 | 6.9 |
| | 1994 | 24.2 | 20.1 | 14.9 | 14.3 | 10.9 | 8.8 | 6.8 |
| | 1999 | 24.9 | 20.9 | 17.1 | 11.0 | 11.4 | 8.3 | 6.5 |
| | 2005-06 | 23.9 | 21.8 | 16.2 | 13.6 | 9.5 | 8.1 | 6.9 |
| | 2010-11 | 21.6 | 19.8 | 18.5 | 14.0 | 11.3 | 7.9 | 6.9 |

These tables provide only a general outline of the population's characteristics, yet it is important to know how the population composition has evolved over time. It has been demonstrated that education is one of the most influential factors driving fertility change. The next section describes the background characteristics in a more detailed manner. Part of the reason for doing so is to evaluate more systematically and assess the quality of the DHS data. Demographers must always look at the results with a cynical eye and always be aware of indicators of the bias that may be introduced by data into the results.

### 3.2.1 Age distribution in Southern Africa

The age structure of women in each country is displayed in Figure 3-1. Note that South Africa and Swaziland are plotted on the same set of axes since they both have one survey that can be analysed. We chose to use the 1998 SADHS, although IIASA used the 2007 Community Survey to eliminate the bias that may arise when harmonising data from different sources. Moreover, DHS data are easy to work with using the *tfr2* STATA macro-command as it was designed to be easily applicable to the DHS standard recode variables (see section 3.3.3).

**Figure 3-1 Country specific percentage age distribution**




The age structure for countries under study shows great similarities from one survey to the next. This is an indication of consistency in sampled women over time with respect to age. There is a change in the population from being dominated by younger women to older women, evident from the age patterns which are seen to be sloping down at younger ages. There are pronounced dips in the 1994 and 1999 ZDHS on 25-29 and 30-34 age groups respectively. The effect is observed on the same cohort of women, which can be explained by the shortage of births in 1967 where most men were recruited and separated from their families to join the armed struggle against the Rhodesian government (Sambisa, 1996). South Africa shows less steep gradient of the age distribution pattern than the other countries in Southern Africa although a similar pattern is observed as well.

### 3.2.2 Educational attainment

We harmonised our DHS data to match IIASA (see section 2.2.1), thereby using the same definitions and principles to define each educational level. Using the procedure summarised in Table 2-3 women are split according to each educational attainment category to generate consistent cross-national age-educational distributions of women. The mapping of the recode variables into six education categories is summarised in Table 3-4.

**Table 3-4 Creating the education variables in Southern Africa**

| *IIASA 2015* | *Lesotho, Namibia, Swaziland and South Africa* | *Zimbabwe* |
|---|---|---|
| No education | $v133=0$ and $v106\leq 1$<br>$v133<1$ and $v106=1$ | $v133=0$ and $v106=0$<br>$v133<1$ and $v106=1$ |
| Incomplete primary | $1\leq v133<6$ and $v106=1$ | $1\leq v133<7$ and $v106=1$ |
| Completed primary | $v133=6$ and $v106=1$<br>$6<v133<9$ and $v106=2$ | $v133=7$ and $v106=1$<br>$7<v133<9$ and $v106=2$ |
| Completed lower secondary | $9\leq v133<12$ and $v106=2$ | $9\leq v133<12$ and $v106=2$ |
| Completed upper secondary | $v133=12$ and $v106=2$<br>$v133<13$ and $v106=3$ | $v133=12$ and $v106=2$<br>$13\leq v133<14$ and $v106=3$ |
| Post-secondary | $v133\geq 13$ and $v106=3$ | $v133\geq 14$ and $v106=3$ |

The education distributions for countries under study are presented in Figure 3-2, Figure 3-3 and Figure 3-4. In these countries where more than one DHS is available, it is observed that the proportions of uneducated women decline consistently over time. In 2014 most women in all age groups had completed primary education in Lesotho, and lower-secondary in Namibia in 2013. The last observed DHS in Zimbabwe indicates that most women have completed lower- or upper-secondary.

However, there are some inconsistencies observed in Lesotho in the cohort of women aged 40-44 in the 2004 survey where a significant drop in the proportion of uneducated women from 13 to 5 percent in 2009. This is a result of a change in definition of the questions in the women's questionnaire. The women's questionnaire for Lesotho 2004, question 109 only asks "what is the highest grade (standard/form/year) you have completed in that level?" By contrast, the equivalent question in the 2009 DHS was modified to be more precise, urging the enumerator to record a '0' for a person who did not complete the first year at the level attended or being attended by the time of the survey.

**Figure 3-2 Educational distribution by age and survey in Lesotho and South Africa**

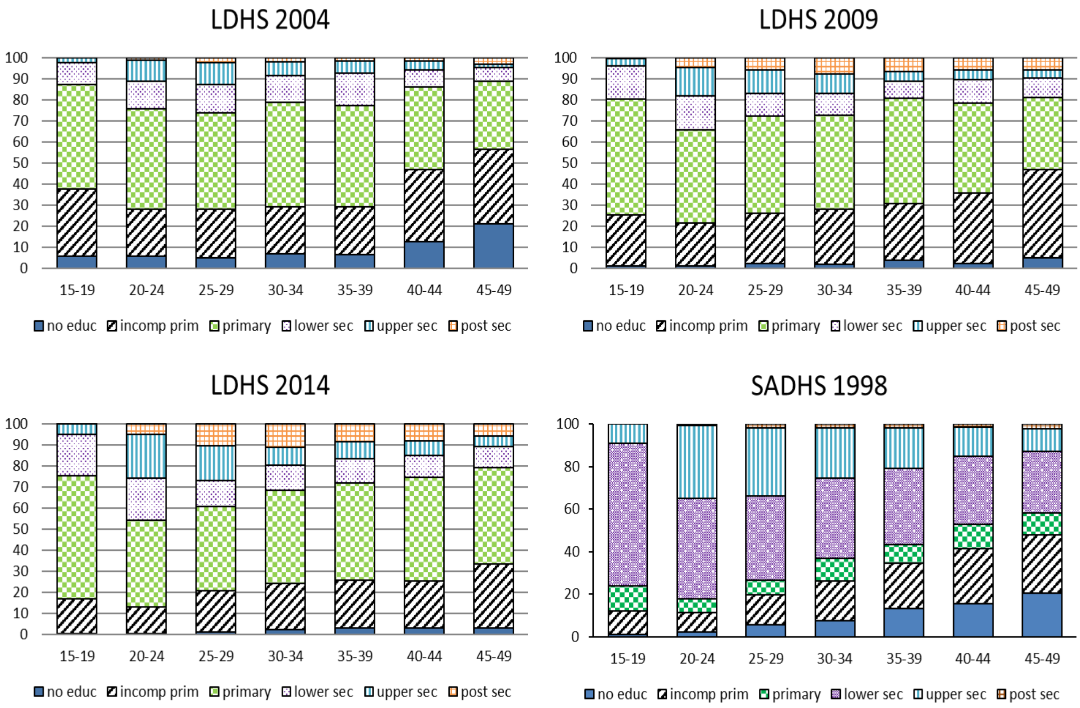



**Figure 3-3 Educational distribution by age and survey in Namibia**

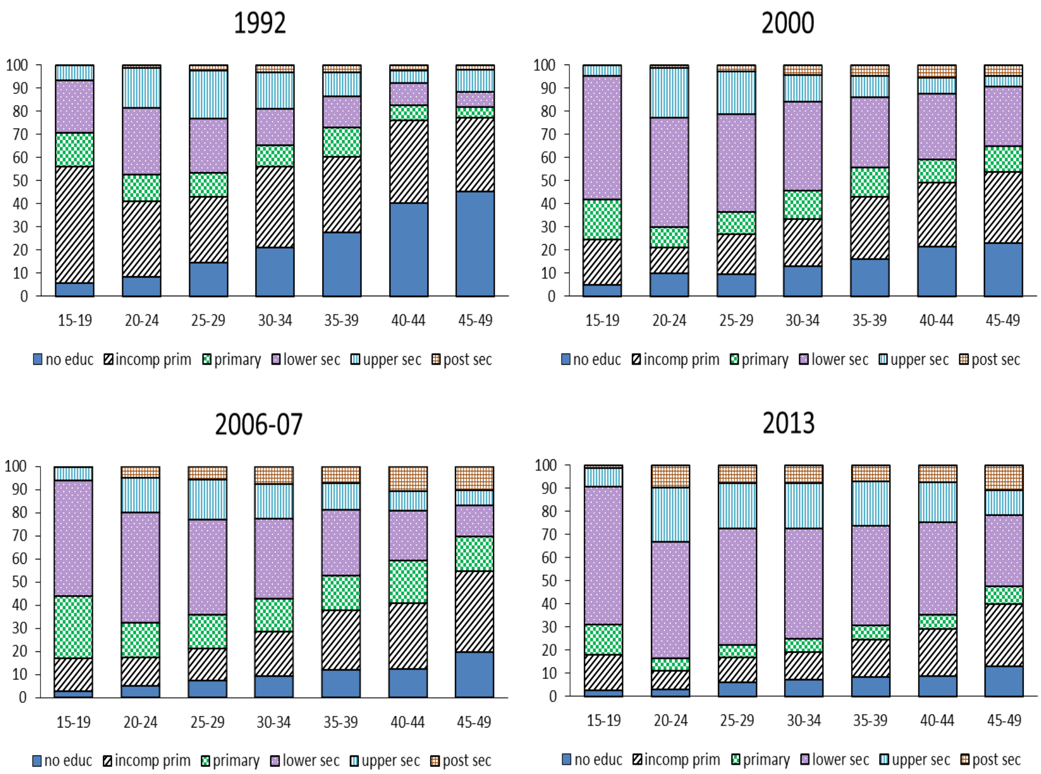
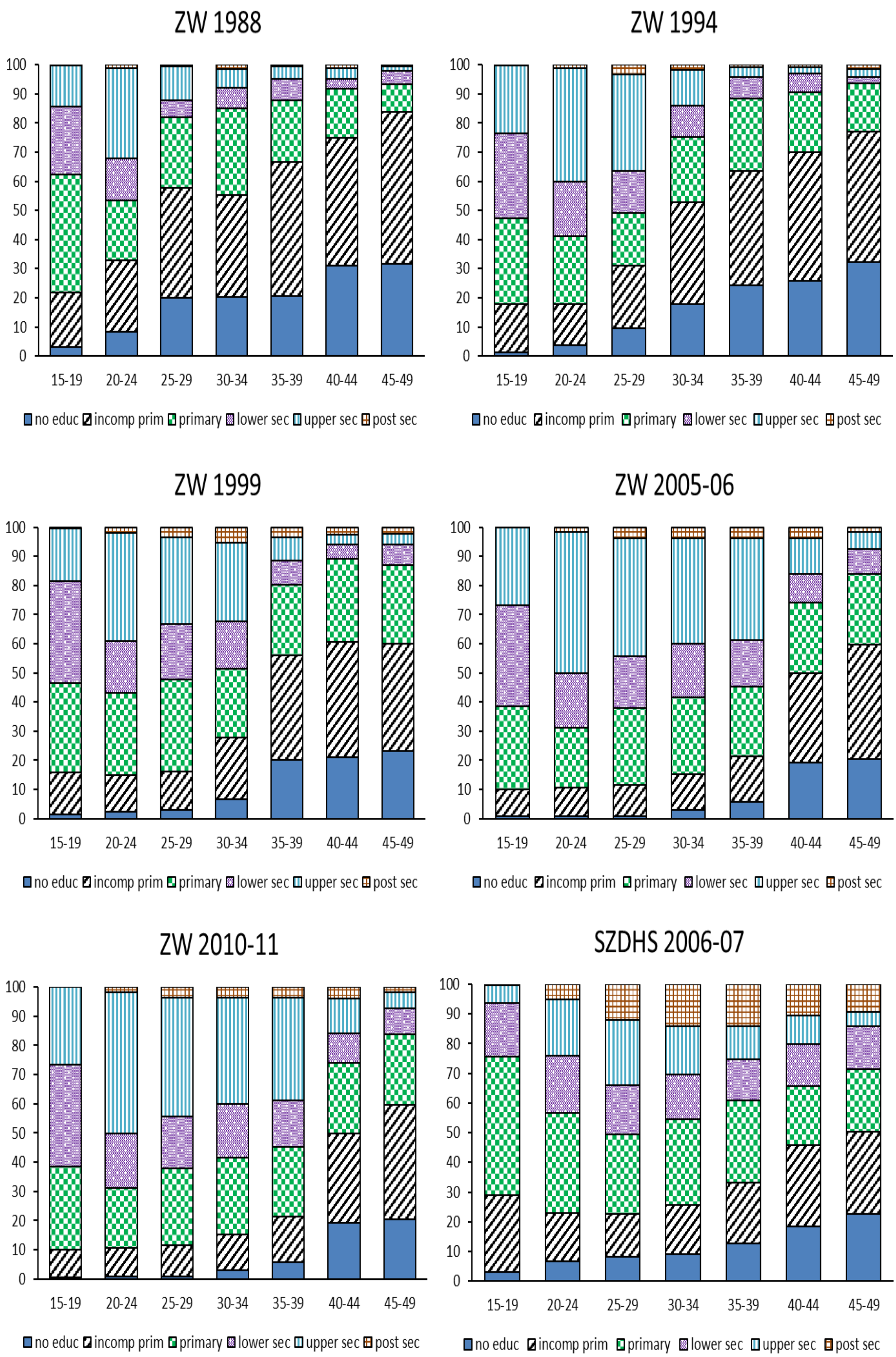


**Figure 3-4 Educational distribution by age and survey in Swaziland and Zimbabwe**



### 3.2.3 Fertility in Southern Africa

Having described the data and the background characteristics, this section investigates and describes the levels of fertility implied by the respective DHSs in each country. We use the standard measures of fertility to describe fertility change in the region: Total Fertility Rates (TFRs) and Age-Specific Fertility Rates (ASFRs). Standardised ASFRs are also used to check the differences in the age composition in each subgroup of the population.

The TFR is a period measure, and hence is affected by the changes in the timing of fertility. For example, the TFR for a three-year and five-year period before the same survey are different, because of how the births are distributed over each period. Generally, the DHS computes fertility rates for the three-year periods before the survey date to avoid unwanted fluctuations in fertility estimates which may arise from relatively small annual births in the survey. However, in the present study we compute fertility rates by five-year age groups and five-yearly period before the survey because we further disaggregated our sample into smaller sub-groups by education levels.

When evaluating the quality of data on recent fertility, the major check which must be done is the assessment of the plausibility of the distribution of ASFRs calculated directly from the data. Plausible fertility distributions are characterised by their uni-modal and concave shapes, being slightly skewed to the right and the ASFRs are close to zero at the extreme of childbearing the age range (45-49 years). This is shown in Figure 3-5 to Figure 3-8, the shape of fertility distribution is reasonably stable in Southern Africa.

**Figure 3-5 ASFRs five years before the survey in Lesotho**

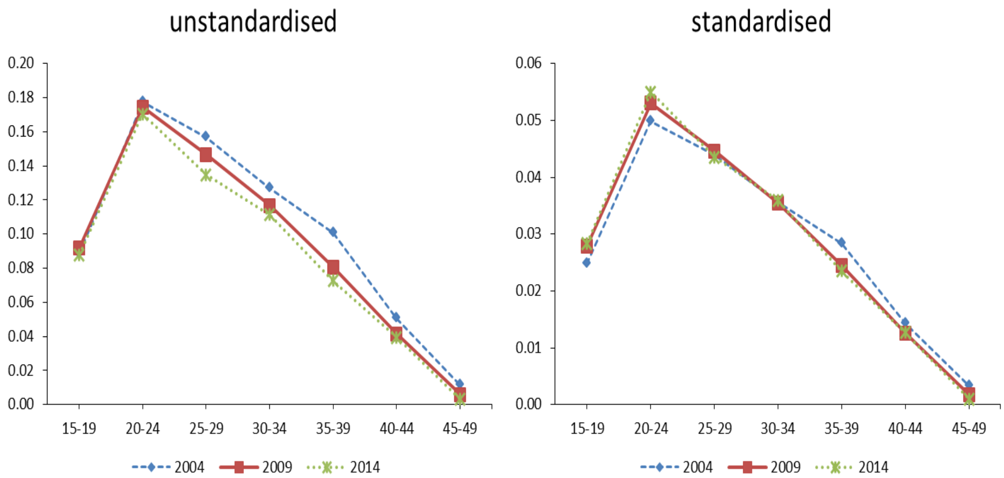


**Figure 3-6 ASFRs five years before the survey in Namibia**

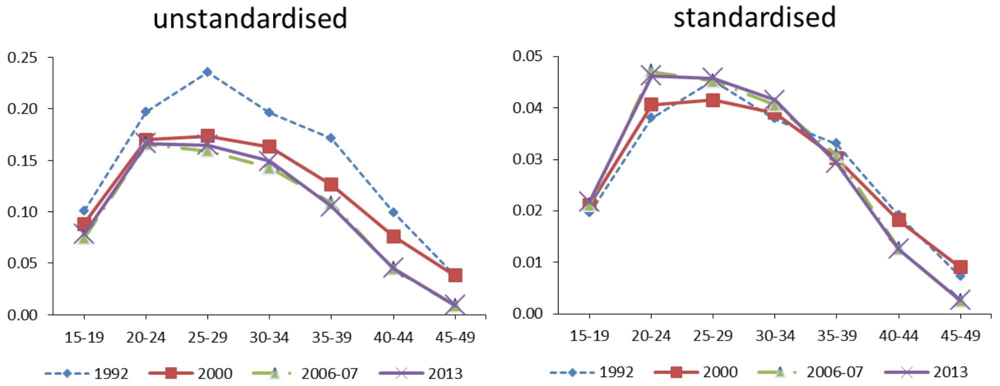



**Figure 3-7 ASFRs five years before the survey in South Africa and Swaziland**

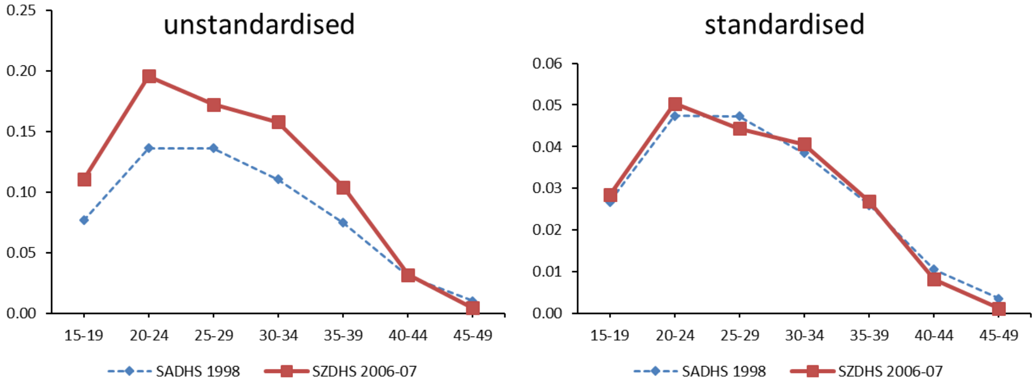
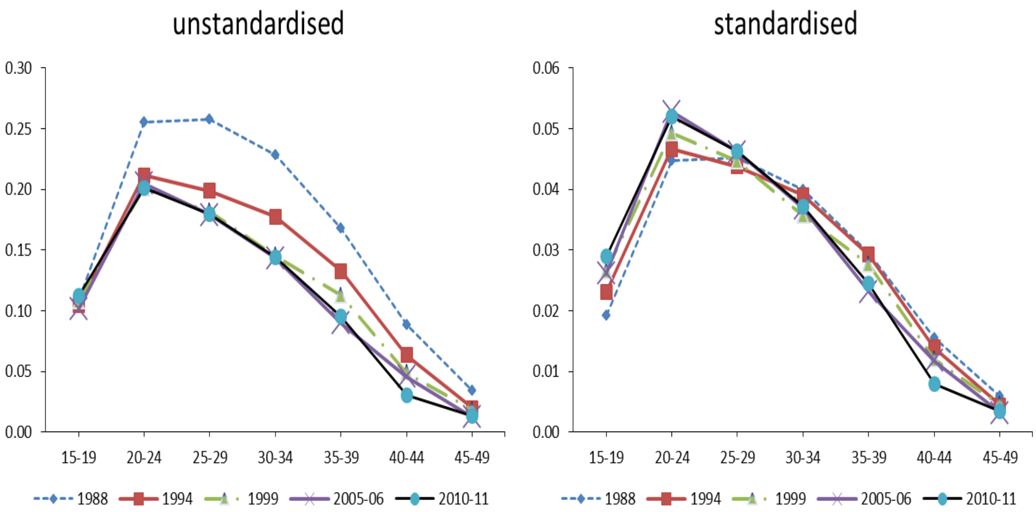



**Figure 3-8 ASFRs five years before the survey in Zimbabwe**




The TFR was back-projected (reconstructed) using the *tfr2* STATA macro command over a period of 15 calendar years prior to each DHS in each country. Birth histories are truncated and fertility rates cannot be estimated for ages and periods corresponding to the oldest woman (for example, ages and periods above the diagonal line in a Lexis diagram). However, the *tfr2* STATA macro command produced by

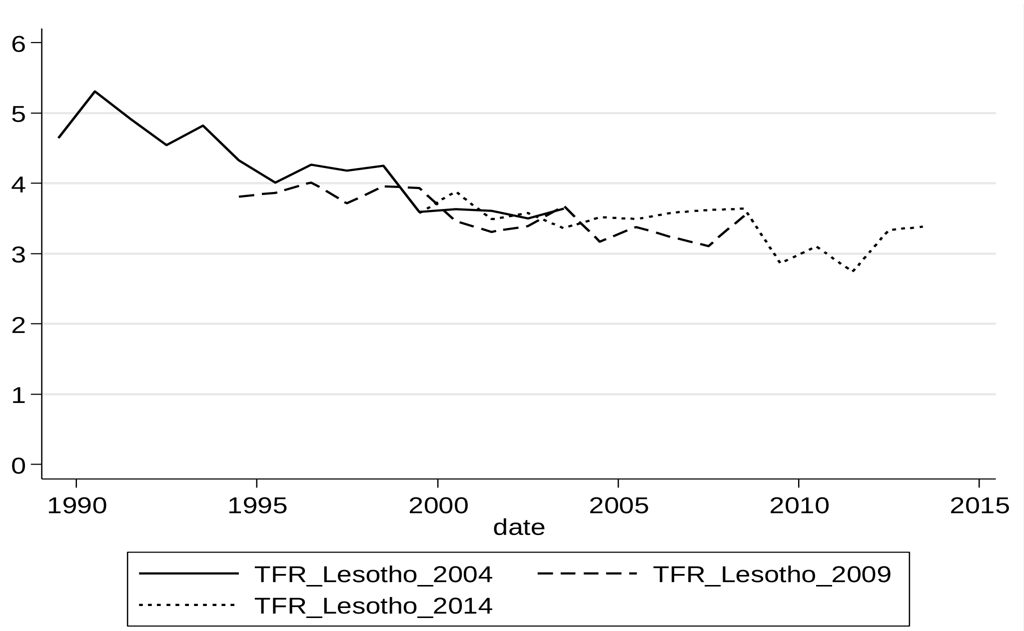
Schoumaker (2013) uses a Poisson regression model with the assumption that the age pattern of fertility remains constant, thereby allowing reconstruction of the TFR for women in the 15-49 age range for periods which we would not ordinarily be able to produce estimates as a result of the age range for which women were selected for inclusion in the DHS.

This approach has several advantages since it offers an important way to evaluate the quality of fertility data. For example, when reconstructing TFR by calendar years a sudden drop in TFR at a cut-off year of a period of inquiry indicates the probable presence of errors in reporting of births. Women may displace births further back in time to avoid additional questions such as, questions about child health and anthropometry module (Moultrie, Sayi and Timæus, 2012, pg 110). This happens especially when women are reporting births which occurred a few years before the survey.

#### *3.2.3.1 Lesotho*

Figure 3-9 depicts a consistent decline in fertility that started in the early to the late 1990s in Lesotho. Despite the visible downward trend in TFR, this figure also illustrates small discrepancies across the surveys, suggesting potential data quality problems in the surveys, possibly reflecting omission and or displacement of births.

**Figure 3-9 TFR (15-49) for 15 calendar years before survey, Lesotho**

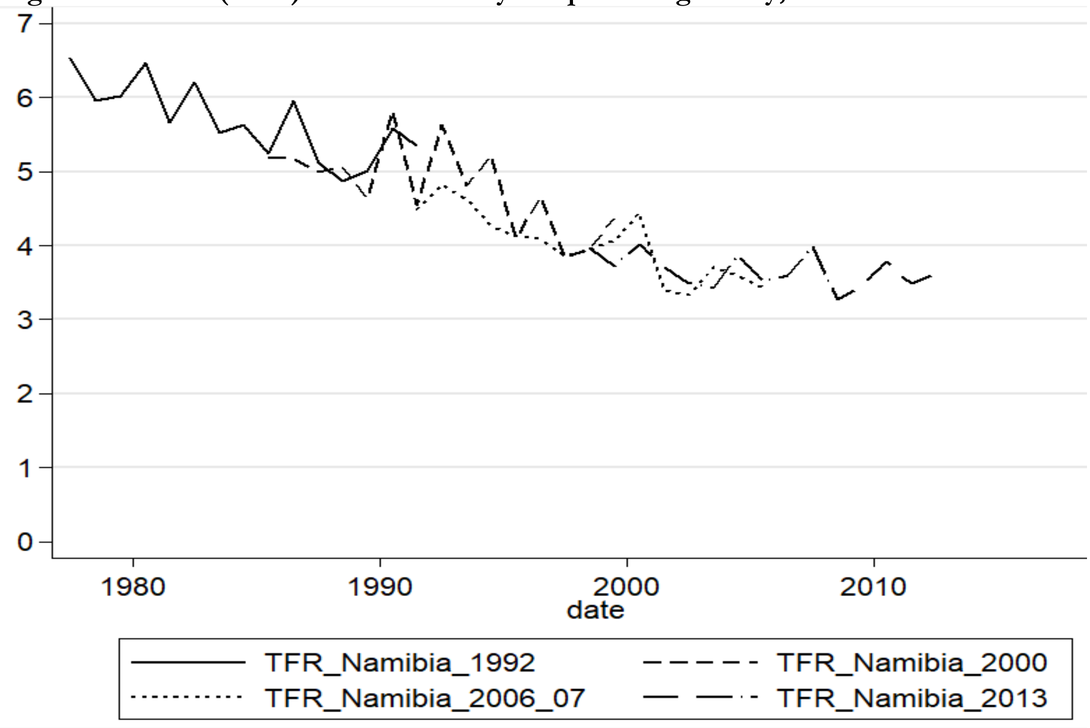



In Lesotho, we observe small discrepancies between recent fertility in the 2009 DHS, and fertility in the same period estimated from the 2014 DHS, suggesting some displacement of births in the 2009 survey. The sudden drop of TFR (from 3.6 to 2.8) at the start of a health module (year 2009) clearly suggests the same problem in the 2014 LDHS.

*3.2.3.2 Namibia*

Figure 3-10 shows the reconstructed fertility trends fifteen years before each DHS in Namibia. Fertility has been declining since the late 1980s through 2006, and fluctuates around 3.5 children per woman afterward until 2013. Therefore, fertility stalled since 2000 in Namibia. There are points in time where different surveys give the same TFRs, which is an indication of consistency that is exhibited in these surveys. There is a sudden drop in TFR in the year 1991 in the 2000 survey and in 2001 on the 2006-07 survey. This sudden drop, though not severe, probably suggests the displacement and/or omission of births in these two surveys. The spikes which are observed possibly reflect digit preference in reporting birth dates.

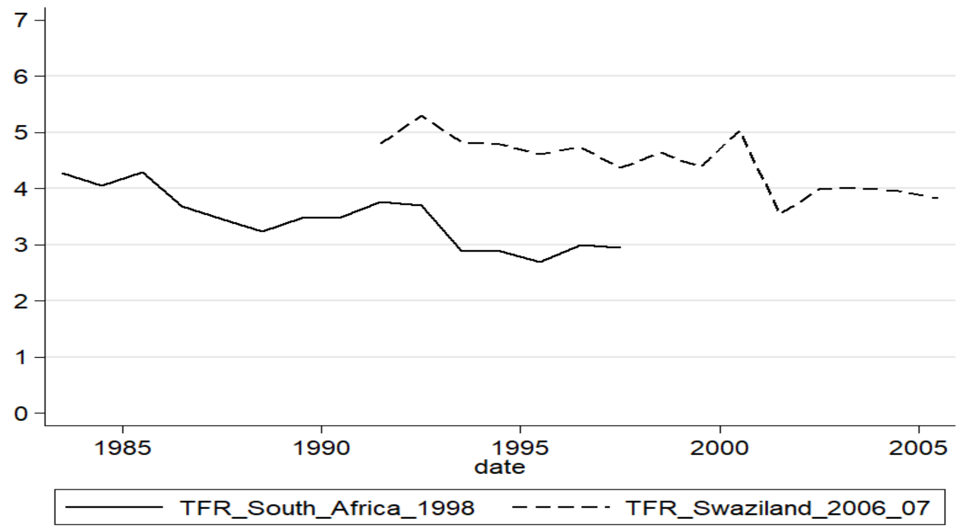
**Figure 3-10 TFR (15-49) for 15 calendar years preceding survey, Namibia**




*3.2.3.3 South Africa and Swaziland*

Figure 3-11 shows that fertility has been falling in both countries. Again, there are sudden declines in fertility in the 'cut-off' year of 2001 in Swaziland and in 1993 in South Africa. This indicates errors caused by women avoiding questions involving anthropometry and child health.

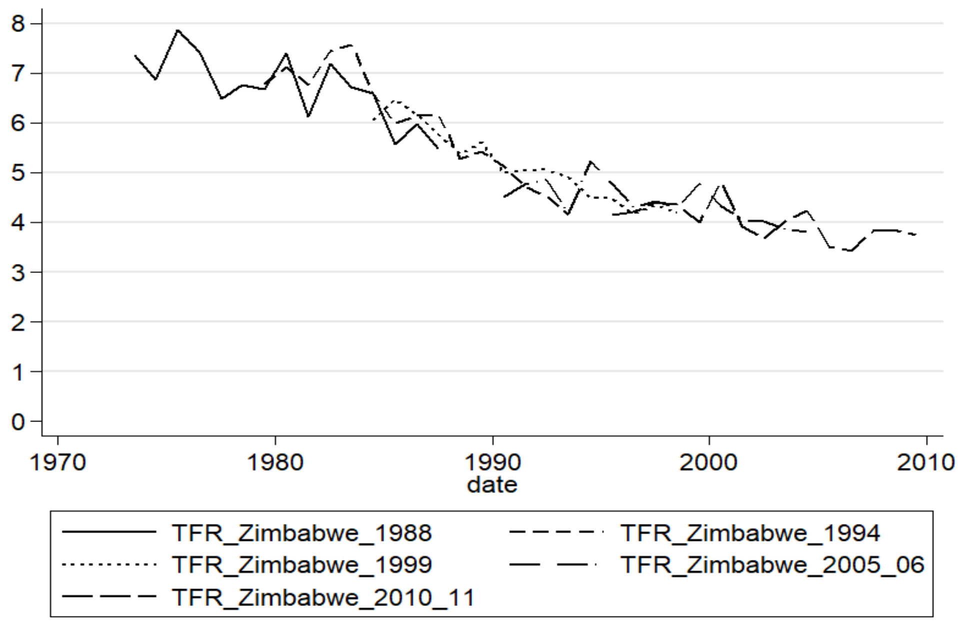
**Figure 3-11 TFR (15-49) for 15 calendar years before survey, SA- and SZDHS**




### *3.2.3.4 Zimbabwe*

Comparing the reconstructed total fertility (TFR) trends from successive surveys (Figure 3-12), there is a clear downward fertility trend in Zimbabwe, which started in the mid-1970s until early 2007, and thereafter it stalled around 3.9 children per woman. The figure also illustrates some minor discrepancies across surveys, suggesting minor data quality issues in some of the surveys. For example, fertility in the 1988 survey seems to be underestimated, possibly reflecting omissions of recent births. Most of the fertility in the same periods estimated by different surveys tally, which is an indication that there are no significant differences in the samples drawn or selected from independent surveys.

**Figure 3-12 TFR (15-49) for 15 calendar years preceding survey, Zimbabwe**



## 3.3 Methods

This section starts by describing the creation of variables before we outline the methods used, then finally we discuss the procedure where a model was applied to fit a general trend for fertility change in the region and estimates generated from this model are compared with the WPP estimates.

### 3.3.1 Creating the education variables

The education variables are generated from each DHS using the two variables in the datasets: (1) education in single years (*v133*); (2) highest education level (*v106*) which is coded from 0 "no-education" to 3 "post-secondary". These recode variables are similar for all countries with DHS data. We then cross-tabulate education in single years and the highest education level gives us a clear indication of how to generate these variables. It must be noted that what we define after recreating variables is not how the DHS would have defined the education categories.

All countries in Southern Africa have similar education structure where a minimum of twelve years is required to complete secondary education except in Zimbabwe where primary education is seven years, lower secondary education is attained by those who have completed two years (Junior Certificate) of secondary education, upper-secondary education is comprised of those who have completed at least four years (Ordinary Level) to a maximum of six years (Advanced level) in secondary, while those with post-secondary education have completed at least one year of tertiary education. This latter class is broad since it encompasses completion of short courses (duration of one year), which leads to an award of a National certificate to more advanced research degrees (such as PhD).

The focus in generating education variables is on years of education completed. Those who have completed at least one year but less than the minimum number of years to complete a specific education level are grouped as incomplete for that level. For example, the incomplete primary comprises those who have completed at least one year but less than seven years of primary education.

### 3.3.2 DHS and IIASA education distributions

We start by comparing the regrouped education distributions by age for the year of the DHS against the IIASA distributions. IIASA distributions are given in five-year periods, for example, the 2010 distribution covers a period from July 2010 to end of June 2015. Therefore, we compare the DHS educational distribution to the IIASA distribution covering the period closest to the date of the DHS. Education composition does not change rapidly over time. Thus, we expect the DHS and IIASA distributions to be very similar.

A detailed comparison is performed by weighting the ASFRs ( $f_{x,y}$ - where '*x*' is the age-group and '*y*' is the education level) by the probability distribution in each age group. The distributions are obtained from IIASA and compared to the DHS. For instance, in the age group 15-19 if we sum the proportions of women in each education level we must get a total of one: let $i_{x,y}^{t_j}$ be the IIASA education proportion of women aged-*x*- and of education category-*y*- at time-*tj*, then $\sum_{y=1}^{6} i_{x,y}^{t_j} = 1$; $d_{x,y}^{t_j}$ is the DHS education proportion of women aged-*x*- and of education category-*y*- at time-*tj*,.

$$\sum_{y=1}^{6} f_{x,y} i_{x,y}^{t_j} = ASFR_x^{i,t_j} \quad (1)$$

$$\sum_{y=1}^{6} f_{x,y} d_{x,y}^{t_j} = ASFR_x^{d,t_j} \quad (2)$$

$$TFR_{national}^{a} = 5 \sum_{x=15-19}^{45-49} ASFR_x^{a}, \{a = i, d\} \quad (3)$$

**Table 3-5 Age 20-24 ASFRs by IIASA educational distributions in Lesotho**

| $y$ | *No education 1* | *Incomplete primary 2* | *Complete primary 3* | *Lower secondary 4* | *Upper secondary 5* | *Post-secondary 6* | *Weighted ASFRs* |
|---|---|---|---|---|---|---|---|
| ASFR | 0.221 | 0.233 | 0.174 | 0.138 | 0.090 | 0.047 | |
| $i_{20-24,y}^{t_j}$ | 0.030 | 0.303 | 0.404 | 0.121 | 0.111 | 0.030 | 0.176 |
| $d_{20-24,y}^{t_j}$ | 0.060 | 0.220 | 0.477 | 0.130 | 0.101 | 0.012 | 0.175 |

The weighted ASFRs are calculated using the formulas shown by equations (1) and (2)

Weighting the ASFRs by education distributions gives the national ASFRs. Table 3 5 presents the national ASFRs (weighted ASFRs) derived from the IIASA and DHS educational distributions for women aged 20-24 years in Lesotho. These are almost the same, which is entirely coincidental given that the DHS and the IIASA educational distributions are different. It must be noted that fertility rates from the LDHS 2004 apply from January 1999 to December 2003, while the IIASA proportions are applicable from July 2000 to June 2005. It is reasonable, however, to compare the 2004 LDHS education proportions against IIASA 2000 education proportions because of the overlapping time intervals and since the two closely cover the same time periods. Summing the weighted ASFRs and multiplying by five will then give us the national TFR (equation 3).

This procedure is then applied to all the countries under consideration and the national ASFRs derived from the IIASA and DHS probability distributions are compared over time at the country level with the ASFRs derived directly from the DHSs without splitting women into different education categories. The task here is to investigate how well the historical education distributions implied IIASA match the DHS or vice versa. Normally these should be the same as they describe the same countries observed in same periods from the past.

### 3.3.3 Computation of ASFRs and TFRs

Computation of fertility rates require events and properly classified exposure by age over the period of investigation. This can be done manually but is rather tedious. Consequently, a more flexible tool is required to do such computations for large sums of women. We therefore use the *tfr2* STATA macro, which is composed of STATA commands that transform birth history data from the DHS into a table of births and exposure, and a Poisson regression model to compute fertility rates, trends and differentials from a table of births and exposure (see Schoumaker, 2013, pg 1110-1112).

The ASFRs and TFRs are then computed using *tfr2* by age and education level five calendar years before each DHS. Fertility differentials by level of education are analysed by *tfr2* using stratified analyses and a time-series trend of TFRs by education level is produced for all the countries. This involves computing fertility rates by categories of the education variable, which offers a quick way of computing fertility rates by education level without assuming proportionality of rates, which is often unrealistic since the age pattern of fertility is not always constant, and it implies that there is no interaction between age and other covariates such as women's education level.

### 3.3.4 Educational composition changes and fixed ASFRs from the most recent DHS

The procedure of this exercise is described in detail in the steps that follow. For each country under consideration we take the fertility rates calculated using the most recent DHS and assume that fertility remains constant for each education level. We then take these ASFRs into the future and weight them by the IIASA education proportions to get the ASFRs nationally, which are then summed to get the national TFR. For instance, we take the ASFR for the 15-19 age-group and education category '1' (no education) and assume that fertility remains constant thereafter. What we are saying is that the compositional change that is happening and the IIASA data gives us the proportion of women in each education category which changes over time. Thus, if in that education category fertility remains the same but the proportion of women in that category is changing then the national TFR will change accordingly.

The national ASFRs for five-year periods are computed by weighting the fixed ASFRs by the education proportions obtained from the IIASA in different education scenarios. Once the national ASFRs are estimated we then sum them and multiply the result by five to get the national TFR for the five-year periods in the future. This is an imaginary hypothetical exercise, which can be viewed as the reference case scenario, in the sense that 'what if?' fertility remains fixed over time in each education category, then what would be the impact of the change in education composition on the overall fertility levels in the region, provided that all other factors remain constant.

To highlight the impact of education compositional change on fertility decline in the region when we vary education levels the national TFRs generated will then be plotted together with the WPP estimates on the same set of axes, and comparisons made. This exercise shows us how the compositional effect influences the change in the overall fertility while holding fertility constant in each education level. For example, if in future we have more women with higher education even if the fertility in each education level remains constant, then the overall fertility will decline as well.

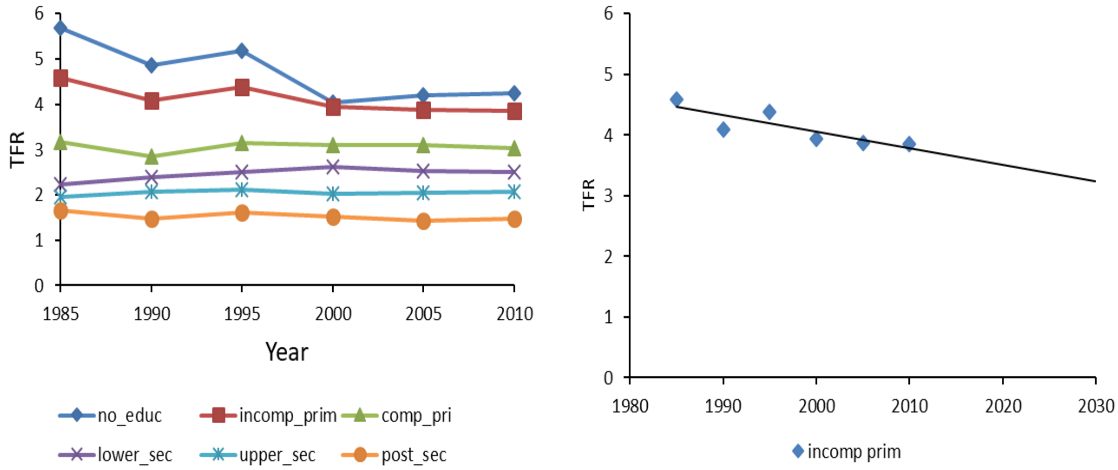
#### 3.3.5 Modelling fertility trends

A model to predict the likely future (2015-2030) fertility estimates (TFRs and ASFRs) by education levels in Southern Africa will be fitted at country-specific levels. To illustrate how this process is done we use Colombia as an example to aid in the exposition of the method. Columbia is a developing country which is near the end of its demographic transition and is chosen because it has similar characteristics to the countries under study. Hence, it is used as an example and it has many DHSs to date that can be used for modelling fertility trends. The idea is to derive and test an approach in one situation, and if it works there, we will test it in other situations.

The first step is to calculate the TFRs by education level five years before each DHS. The TFRs for each education category are plotted against time, which present a general trend of fertility by education level over time. This plot leads us to the formulation and implementation of a model to project fertility. Several steps which are required to fit a model are presented in the paragraphs that follow.

The first step in modelling fertility trends is to choose a baseline category. In our modelling exercise the incomplete primary education group is chosen as the baseline category since the sample size in this category is relatively high in many developing countries and large samples generally yield robust estimates. In Colombia, there are six DHSs up to date and the TFRs over time are plotted in the left panel of Figure 3-13. The benchmark (or baseline) category is shown on the right-hand panel and a linear trend is observed. We then extrapolate the baseline TFRs into the future from 2015 to 2030.

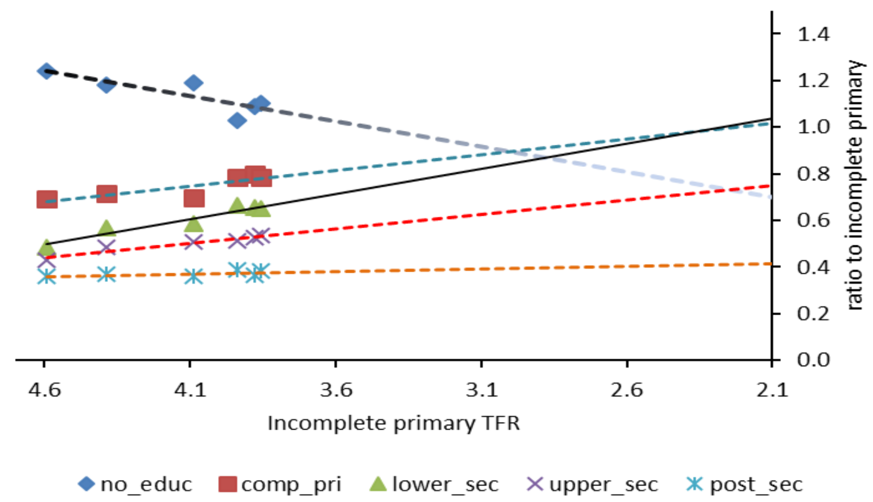
**Figure 3-13 TFRs by education and trend line fit for incomplete-primary, Colombia**




A linear model will be used to project fertility in this education category, since we are considering a short projection period (2015-2030) this model is reasonable. It must be noted that this may not necessarily be robust although it seems reasonable for short-projection intervals. We must exclude some DHS data points in this exercise if, for example, fertility in the baseline category changed rapidly in the earliest surveys and gradually in recent surveys to avoid the more distant data points characterised by rapid fertility decline to influence recent fertility change.

The second step is to express fertility in other educational categories relative to the baseline category for each DHS. The next step is to plot these ratios against the baseline TFRs; this may be used as a diagnostic plot to check how these ratios are related to the baseline fertility. It must be noted that there is no theory that tells us that the ratios should be linearly related to the baseline TFRs, but from the Colombian plot (see Figure 3-14) it was seen that these are approximately linear. Therefore, a linear relationship between the ratios and the baseline TFRs is assumed to apply to the countries under study. The assumption is reasonable for short projection periods, such as the one we are considering in the study. However, this is problematic for long-term projections and further developments must be considered if this model is to be applied. We then extrapolate these ratios forward in time (from 2015-2030) using all the points since the assumption of linearity is satisfied. From the graph we observe significant changes in the ratio of women without education, this seems to be problematic and considering that this category has relatively small numbers of women compared to other education categories. The contribution from this category must be interpreted with caution as it may imply implausible fertility rates (crossing-over of ratios beyond TFR of 3.1).

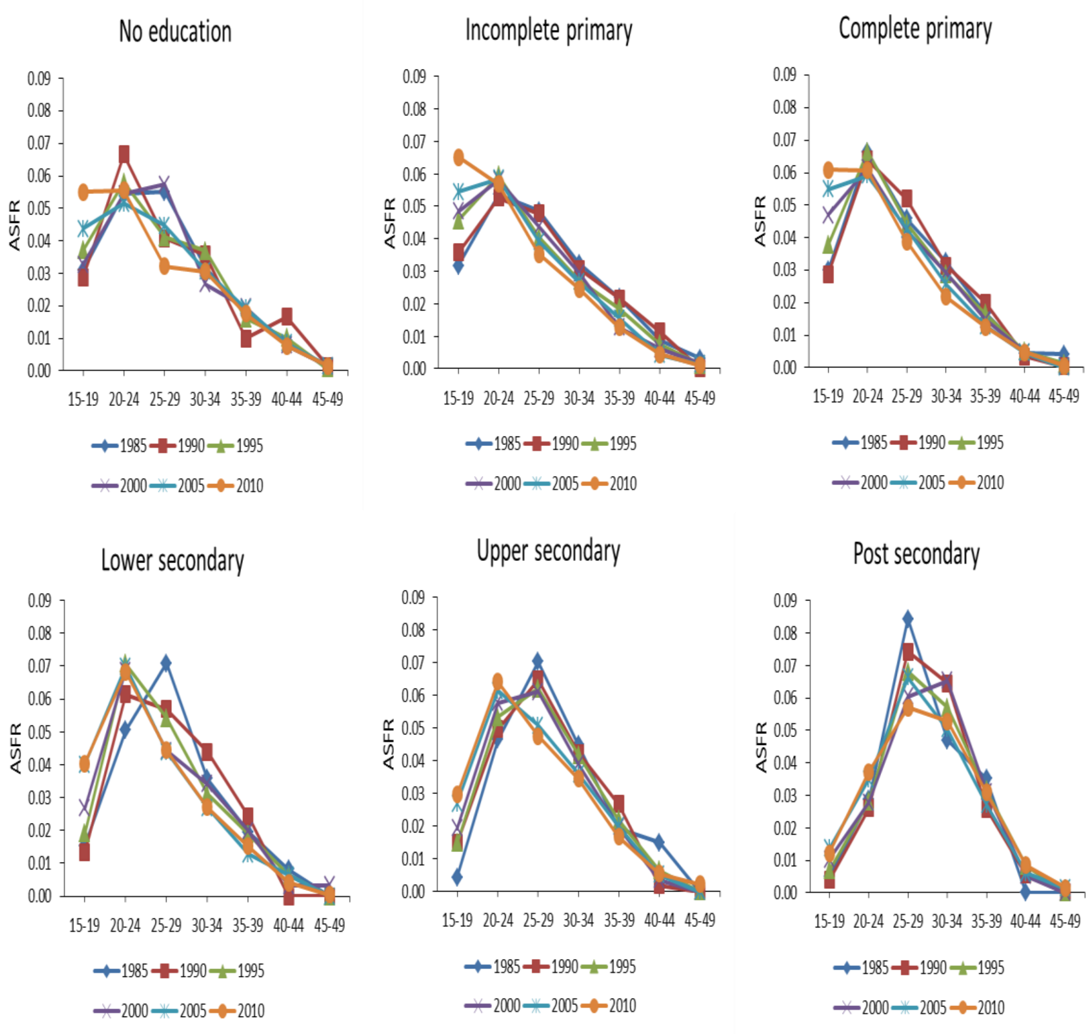
**Figure 3-14 Baseline TFRs against ratios of education levels, Colombia**




The projected ratios are then multiplied by the projected baseline TFRs corresponding to the projection interval of inquiry. Thus, obtaining the TFRs in the other five education categories. We now have TFRs by education level from 2015 to 2030. The corresponding ASFRs for future years are then calculated as follows.

We use the standardised fertility rates to calculate the ASFRs that correspond to the projected TFRs. The fertility-structure does not change rapidly over time; thus, we use these to find the ASFRs by education level. Each TFR by education category is distributed according to the fertility structure to attain the projected ASFRs. The age pattern of fertility does not seem to change much over time in Colombia as shown in Figure 3-15. Although small discrepancies are observed in younger women we chose to use the average of the distributions as the one applicable to the future. This has turned out to be a good approximation in Colombia, and the same technique may be adapted to the respective countries in Southern Africa. However, if this age-pattern of fertility is changing rapidly from the earliest surveys to the most recent, one would need to allow for this.

**Figure 3-15 Standardised ASFRs, Colombia DHSs**




Finally, once these projections are presented we must check for plausibility against what is currently known about fertility. The different education scenarios will be used to check how national fertility is likely to change if educational policies were to change in a manner described by the IIASA. The national fertility rates for different education scenarios are calculated by weighting the projected ASFRs by the education attainment proportions from IIASA under different education policies. The national TFRs are then calculated from these national ASFRs. These are then compared with the WPP fertility estimates. The validation exercise was performed in Colombia, where the model was fitted using the first four DHSs to predict the TFRs for the fifth and six surveys. The difference between the observed and projected values were less than 5 percent for both surveys. Thus, the two were not significantly different from each other. This exercise was applied to other countries with sufficient data to further validate the model.

### 3.4 Conclusion

We began by describing and giving background information of the data used in this dissertation and where they were extracted. DHS data was used for analyses, and characteristics and country-specific evaluations of the data presented. Finally, we presented the methodology from the creation of education variables to fitting a model to project fertility from 2015 to 2030. In order to explain the model and show that it is applicable, we fitted the model to Colombia.

# 4 RESULTS

This chapter presents the results of comparisons of the national fertility rates weighted by Demographic and Health Survey (DHS) and International Institute for Applied Systems Analysis (IIASA) proportions (section 4.1); fertility rates by education level (section 4.2); an analysis of change in education composition with fixed fertility rates (section 4.3); and finally (in section 4.4) we present the results of modelling of fertility trends in the region. These results are derived by applying the methods described in the previous chapter.

## 4.1 Comparison of the DHS and IIASA distributions

The education variables split women into the six education categories as described in section 3.3. From these we computed the proportion of women by education level at the date of the survey for all the five-year age groups (15-49). The educational attainment distributions were computed for all the DHSs and compared with the IIASA educational attainment distributions; explanations of the procedure are given below.

Figure 4-1 shows a panel for each age group of the Zimbabwean educational proportions for the 2010-11 DHS, these are plotted on the same set of axes with the 2010 IIASA educational proportions for comparisons. There are strong similarities in educational attainment distribution in all the age groups in general between the DHS and the IIASA, which is as expected, considering how the education variables were derived from the DHS data. However, minor differences were observed in three age groups (15-19, 40-44 and 45-49). One reason for such minor differences is that these education attainment proportions do not refer to exactly one point in time, but different times which are close to each other. Similar patterns between the DHS and the IIASA educational attainment distributions were observed in all the other countries considered, and presented in Appendix A.

National ASFRs were then calculated from the DHS in each country and the corresponding weighted national ASFRs by DHS and IIASA educational attainment distributions computed and the results presented for Lesotho, Namibia, South Africa, Swaziland and Zimbabwe. In each case, the national ASFRs implied by the education-specific ASFRs weighted by the IIASA and the DHS distributions are virtually identical, as shown in Appendix B.

**Figure 4-1 Educational distribution, Zimbabwe (DHS 2010-11 and IIASA 2010)**

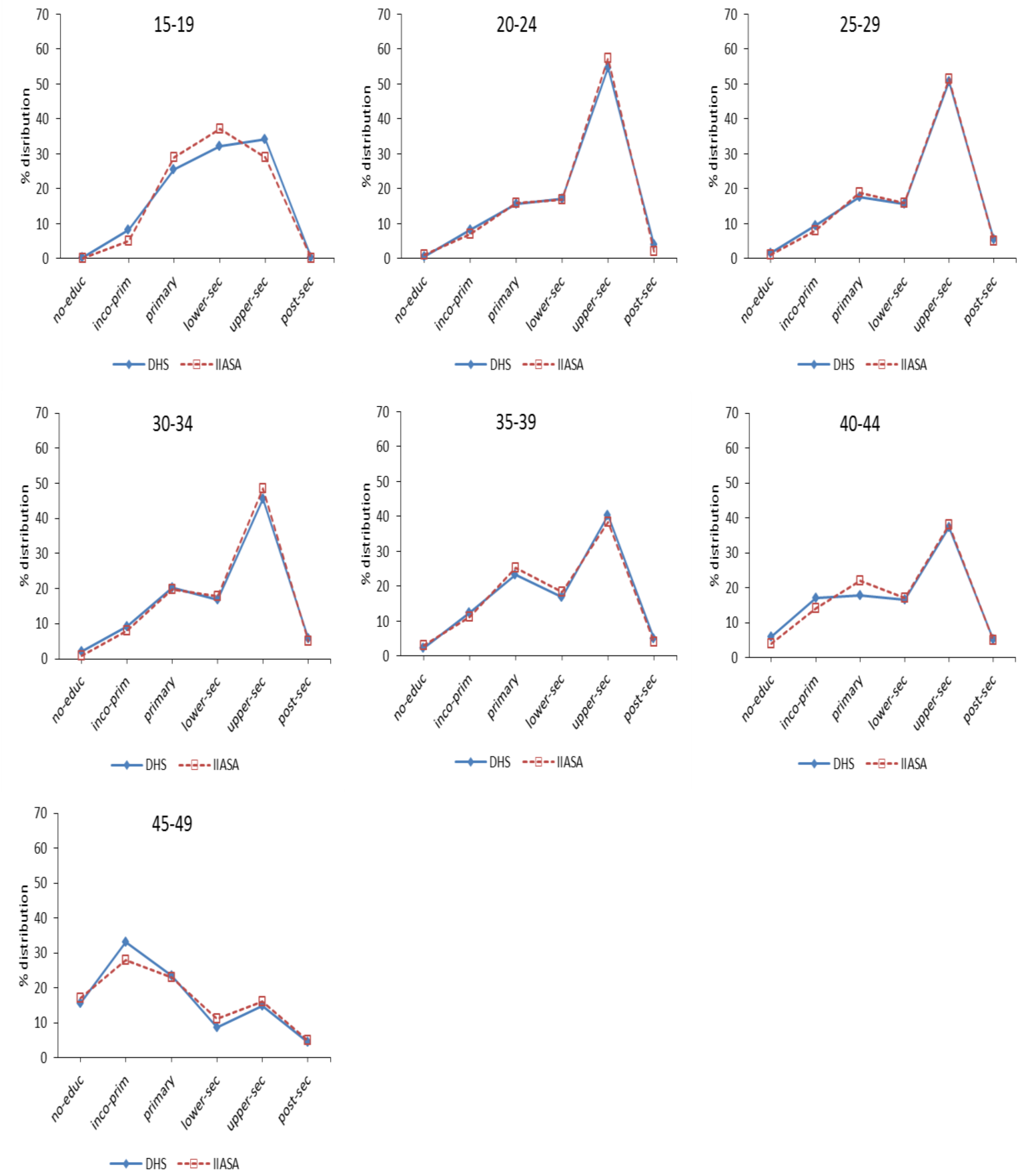



Table 4-1 presents the total fertility rates (TFRs) for each country in Southern Africa. The percentage difference (DHS to IIASA) is less than 5 percent except for the latest Lesotho DHS, which is 5.9 percent greater than the one implied by the IIASA distribution of education status. The national TFRs for all the countries under consideration are not significantly different from each other. This is consistent with the similarities in educational attainment distributions between the DHS and the IIASA within the projection interval.

**Table 4-1 TFRs DHS vs IIASA for each country and by survey date**

| *Country* | *DHS year* | *DHS* | *Weighted by DHS proportions* | *Weighted by IIASA proportions* |
|---|---|---|---|---|
| | 2004 | 3.6 | 3.6 | 3.6 |
| Lesotho | 2009 | 3.2 | 3.3 | 3.4 |
| | 2014 | 3.0 | 3.4 | 3.2 |
| | 1992 | 5.1 | 5.2 | 5.0 |
| Namibia | 2000 | 4.2 | 4.2 | 4.1 |
| | 2006-07 | 3.5 | 3.5 | 3.5 |
| | 2013 | 3.5 | 3.6 | 3.7 |
| South Africa | 1998 | 2.9 | 2.9 | 3.0 |
| Swaziland | 2006-07 | 3.8 | 3.9 | 3.9 |
| Zimbabwe | 1988 | 6.0 | 5.9 | 6.2 |
| | 1994 | 4.7 | 4.7 | 4.8 |
| | 1999 | 4.2 | 4.2 | 4.3 |
| | 2005-06 | 3.9 | 3.9 | 3.9 |
| | 2010-11 | 3.6 | 3.7 | 3.6 |

## 4.2 Fertility trends by education levels

Figure 4-2 shows the percentage distribution by education level for the three countries in the region with at least two DHSs (Lesotho, Namibia and Zimbabwe) in order to examine secular trends in fertility by education level. We need multiple DHS surveys for this analysis rather than just back-casting a single survey to show how fertility has changed over time between two independent points in time. The proportion of women with no education and incomplete primary education has declined consistently in all three countries from the earliest to the latest DHS, while the proportion of women with upper secondary and post-secondary education increased monotonically over time.

The two latest DHSs for each country indicate that in Lesotho, most women have completed primary education, while in Namibia the majority are those with lower secondary education, and Zimbabwe has its majority of women having completed upper secondary education.

Figure 4-3 shows the total fertility rates (TFRs) for the five Southern African countries by education level.  Fertility is higher among women with no education (blue lines) and incomplete primary education (brown lines) in the region. As expected, fertility is lower among educated relative to uneducated women.

**Figure 4-2    Percentage distribution by level of education, women aged 15-49**

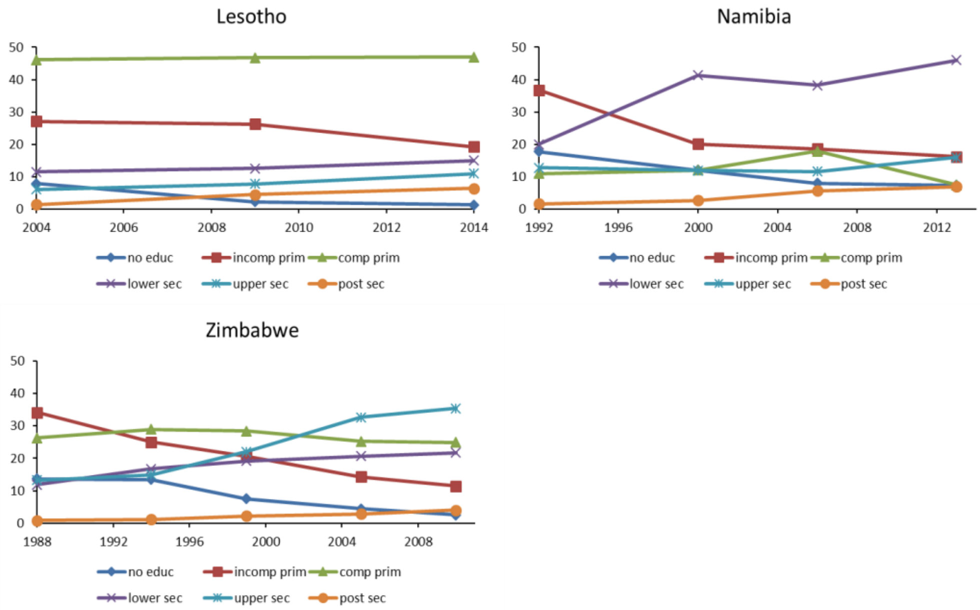
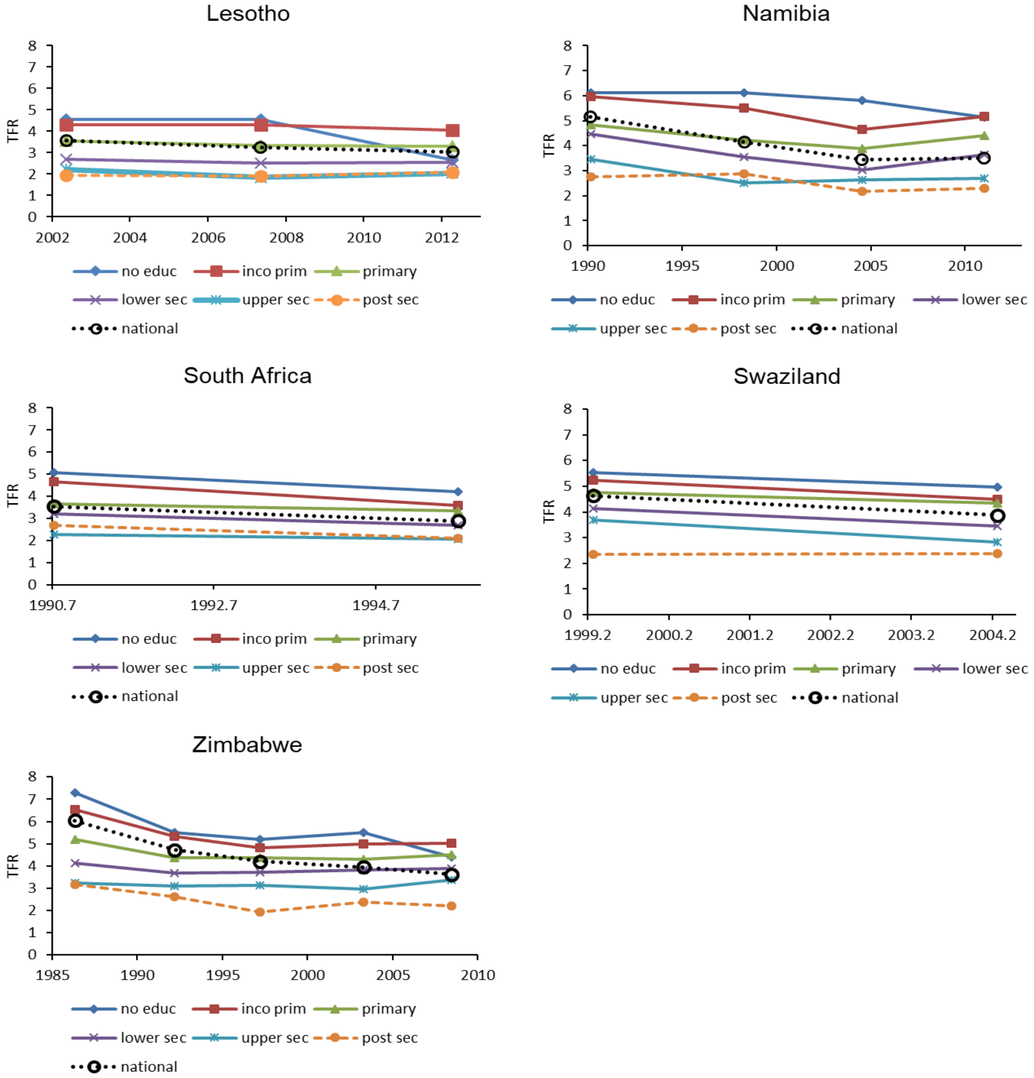


**Figure 4-3 TFR by level education and at national level**




There is an increase in the proportion of women in higher education levels (Figure 4-2), which is consistent with the national fertility decline (Figure 4-3) across all countries. The results at country level shown in Figure 4-3 are discussed in the sections that follow.

### 4.2.1 Lesotho

Lesotho fertility trends by education level have changed steadily over time. Total fertility among women with incomplete primary declined from 4.4 to 4.0 children per woman from 2002 to 2012 and a small decline in fertility was also observed across other education levels.

Fertility declined sharply among uneducated women (from 4.5 to 2.5 children per woman). This is an unexpected result in such a short period, since fertility changes gradually over time, which raises concerns and needs further investigation. Looking closely into the data we discovered that only 89 women in the 2014 LDHS were reported to be uneducated from a sample of 6 621 women. The smaller the sample size, the fewer the events reported, which in turn means there will be large confidence intervals around the estimates derived from such small numbers. Therefore, total fertility for uneducated women is unreliable for these reasons as well as discussed in Chapter 3.

Most women in Lesotho have completed primary education (Figure 4-2) and hence the national TFR is increasingly weighted to women in this education category. The national TFR declined gradually over time from 3.6 to 3.0 children per woman between 2002 and 2014.

### 4.2.2 Namibia

Namibia experienced a rapid increase from the late 1990s in the proportion of women with lower secondary education (Figure 4-2), followed by upper secondary and post-secondary education. At lower education levels (no education and incomplete primary) the proportions of women declined steadily. The national TFR in the 1990s to the early 2000s was heavily weighted to women who have completed primary education. From 2005 onwards the national TFR was concentrated among women who have completed lower secondary education. The national TFR declined monotonically prior to 2005, and stalled thereafter.

### 4.2.3 South Africa and Swaziland

As expected, fertility is higher among less educated women as compared to those with the most education. In South Africa the national TFR was 2.9 children per woman in 1995 and in Swaziland it was 3.9 children per woman in 2004. Considering the two different points in time, South African fertility is different from the other countries in the region since it was lower in the late 1990s. None of the countries in the region have reached these low levels currently.

### 4.2.4 Zimbabwe

In Zimbabwe, fertility declined substantially between the 1988 and the 1994 DHS in all education categories. It declined from 7.8 to 5.8 children per woman for women without education, and for those with incomplete primary education from 6.9 to 5.5 children per woman. Between the 1994 and 1999 DHS we observed a smaller decline in fertility across all education groups, and then from 1999 onwards, as in Namibia, fertility appears to have stalled in all education categories. However, the continued decline in the national TFR (Figure 4-3) reflects the changing distribution of women across educational categories (Figure 4-2).

Fertility among women with post-secondary education has remained the lowest of all in the education category. On the other hand, there is a sudden drop in TFR among uneducated women, which is an unreliable estimate possibly caused by smaller sample size in this category.

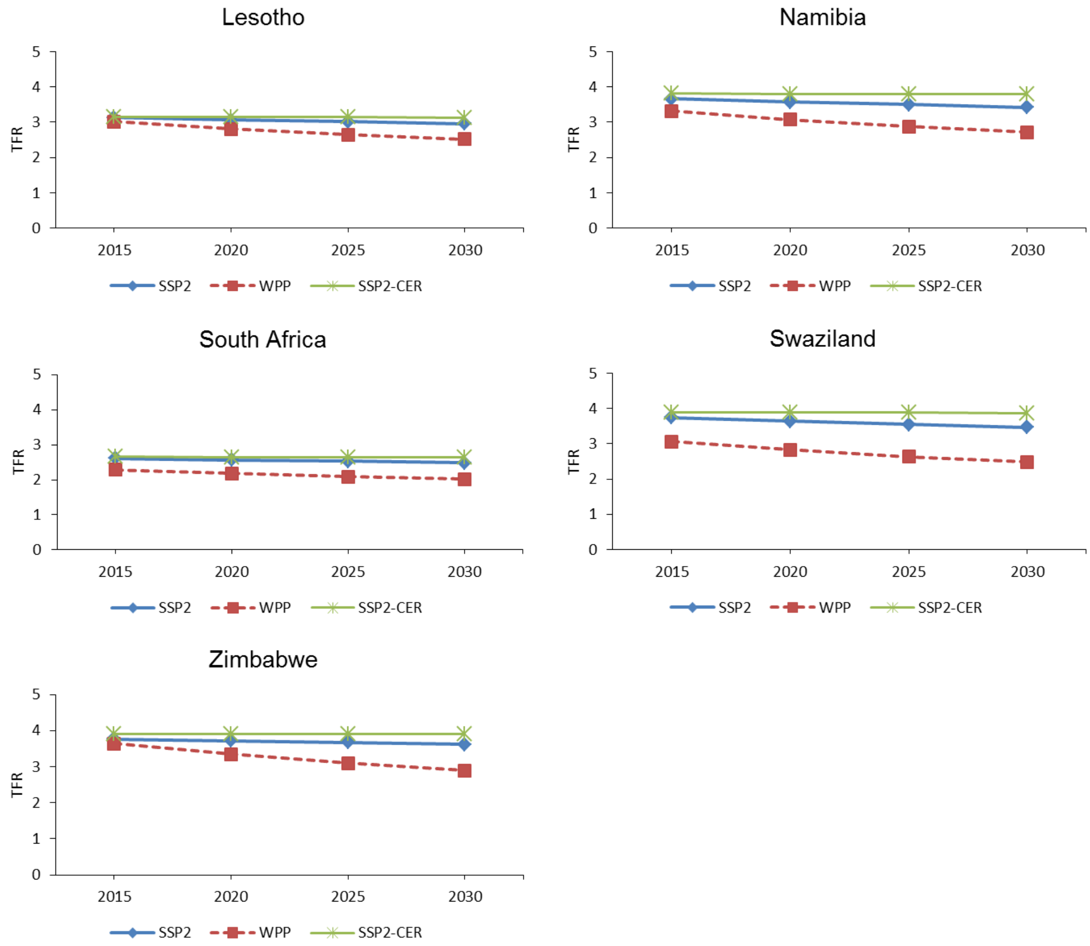
### 4.3 Effects of education composition change

Fertility will not remain constant within each education category over time. However, assuming constant education-specific fertility rates from the last survey allows us to investigate what would happen to the national total fertility when education distributions vary over time. Consequently, this indicates how compositional change effects the national fertility trends in Southern Africa.

The results of the projected national TFRs when we fix ASFRs by education level at the date of the last DHS are shown in Figure 4-4 for all countries under consideration. These are plotted on the same set of axes with the WPP projected estimates from 2015 to 2030 to demonstrate how much fertility decline would be brought about by simple compositional changes in education.

The education scenarios result in different fertility trends caused by education policies implemented in each country. As shown, if fertility of each education level was to remain constant until 2030, there will be a gradual fertility decline observed in all countries in the region and if education was to change in the manner that the SSP2 scenario was met or satisfied, then the overall fertility will decline faster as compared to the SSP2-CER case scenario.

**Figure 4-4 Projected national TFRs and the WPP estimates**




Fertility is expected to behave as shown in Figure 4-4, with fertility declining faster under the SSP2 scenario compared to the SSP2-CER. This is consistent in all countries across the region because SSP2-CER does not change the composition as much. However, increasing differences were observed from 2020 to 2030 as the WPP TFR estimates depart from these scenarios and drops to lower levels, emphasising that the compositional change alone cannot account for the overall fertility decline because the WPP assumes that there will be fertility change. Therefore, a model to explain some of the changes likely to contribute to the fertility decline in conjunction with changes in education composition is used in the section that follows.

## 4.4 Modelling fertility trends

Modelling of fertility trends is based on the methodology set out in section 3.3. Here we show how the model was fitted to each country under study, and present the diagnostic checks performed to fit the model. Finally, we present the modelled TFRs using the two education scenarios (SSP2 and SSP2-CER) and compare these with the WPP projected fertility estimates for each country. The comparison of our TFR estimates with the WPP is to validate our approach for projecting fertility, while incorporating education composition changes.

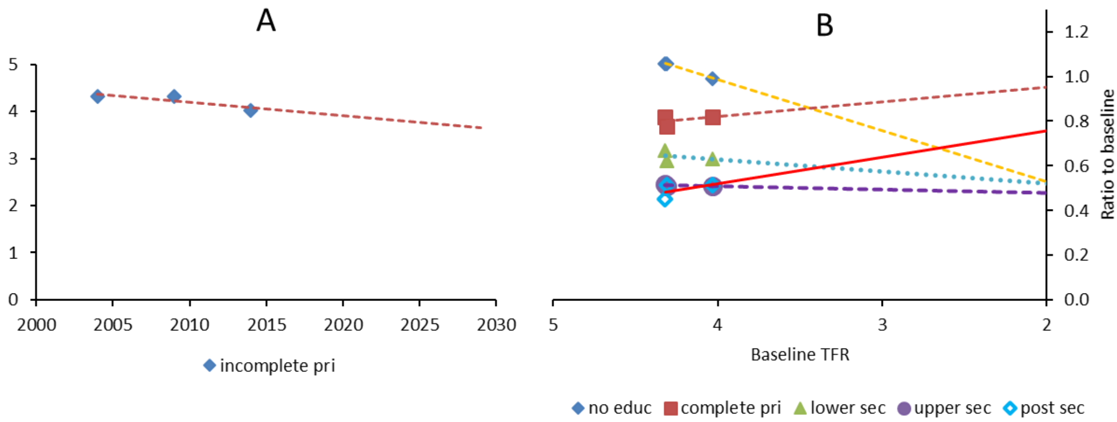
### 4.4.1 Lesotho fertility trends

Using the procedure outlined in the methodology, we first plotted the TFRs for the baseline education category against time and the ratios of other education categories to the baseline to check if our data satisfies the assumptions for the model to be applicable. The major assumption that should be satisfied is that the baseline TFRs and the ratios to the baseline TFRs should be linearly related.

The TFRs for the baseline education categories are plotted (panel A of Figure 4-5) against time using the three LDHSs (2004, 2009 and 2014). From these data points a linear trend line was fitted. All the points were seen to be linearly related, thus we extrapolated the total fertility for the baseline education category forward in time to 2030. After the computation of the ratios for the respective education-specific TFRs to the baseline TFRs, a diagnostic plot follows to check if the linearity assumption holds. The plot of these ratios is shown in panel B of Figure 4-5. It was also satisfactory, as shown by the graph that these ratios are linearly related to the baseline TFRs. Therefore, we used the extrapolated baseline TFRs and these ratios to predict the future fertility rates for the other education categories. At this point the TFRs by education level were projected from 2015 to 2030.

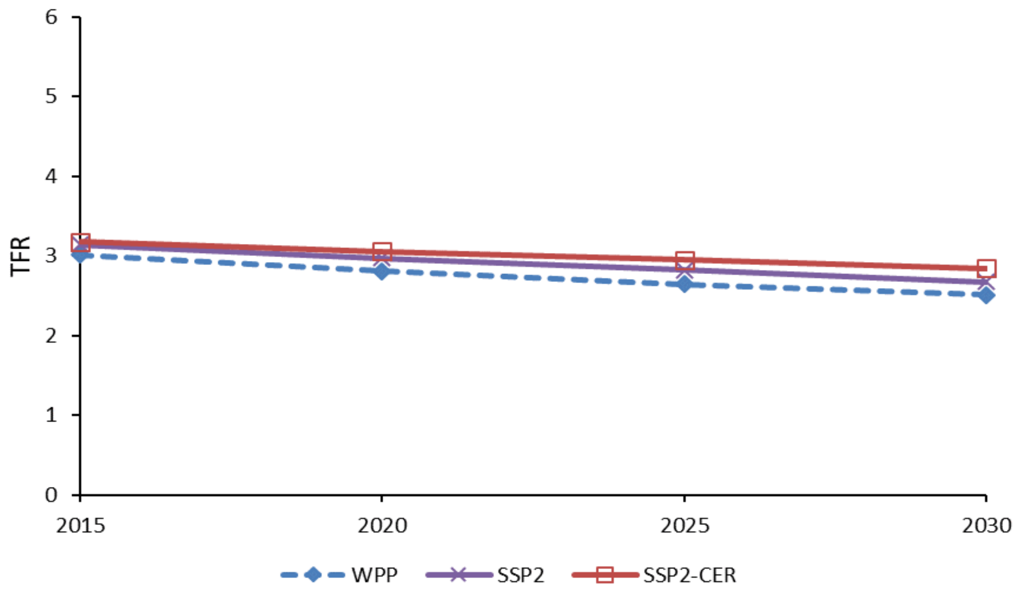
Age-specific fertility rates (ASFRs) by education level were then derived using the average of the standardised ASFRs of the three LDHs. Finally, the national ASFRs were calculated by weighting these age-education-specific fertility rates by the two women's education distribution scenarios adopted from IIASA.

**Figure 4-5 Baseline TFRs and ratios to baseline TFRs, Lesotho**



The results of the national TFRs are shown in Figure 4-6 together with the WPP TFR estimates. If education is to expand in an SSP2 manner the TFRs for each five-year period will differ from the WPP by less than 3 percent from 2015 to 2030. Therefore, the TFRs from SSP2 scenario are similar to those implied by the WPP estimates. Otherwise, if education is to expand in a way that the SSP2-CER education scenario is satisfied then the projected fertility rates will be slightly greater than the WPP projected fertility rates relative to the SSP2 case scenario. There is a 6 percent difference between the WPP and the SSP2-CER TFRs from 2020 to 2025, which decreases to 4 percent in 2030. The percentage difference between the WPP and the fitted TFRs is small considering the projection period of 15 years. Therefore, we can conclude that there is consistency between the WPP fertility estimates and those fertility estimates fitted using these sensible education scenarios.

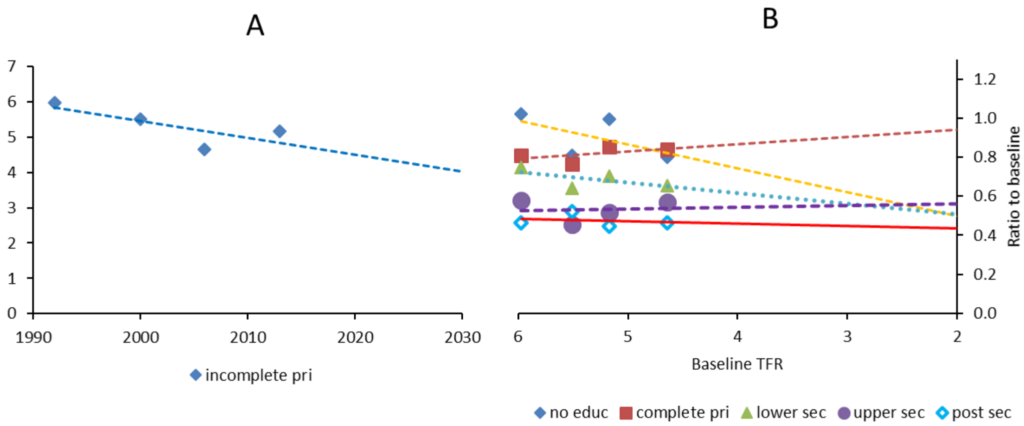
**Figure 4-6 Projected national TFR by education and WPP estimates in Lesotho**




### 4.4.2 Namibia fertility trends

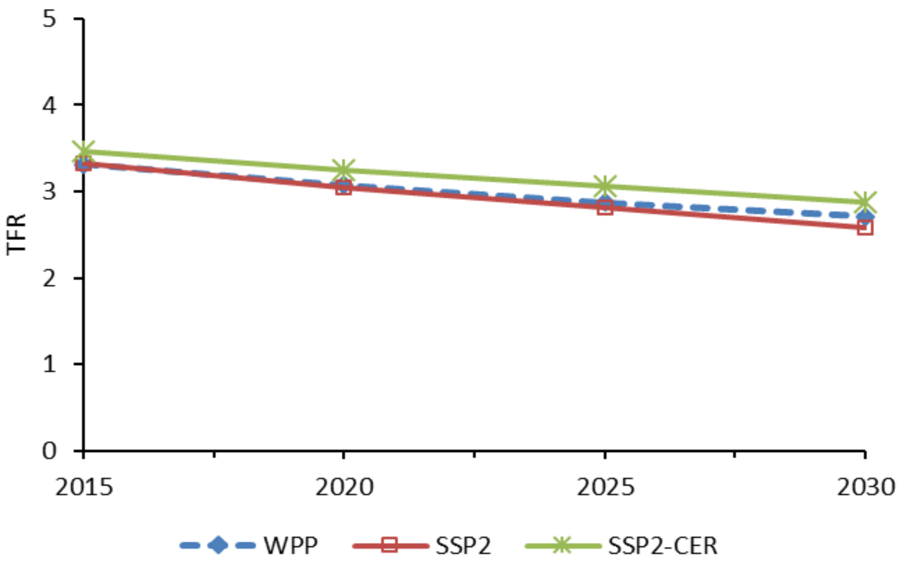
The model was fitted using the four Namibian DHSs (1992, 2000, 2006-07 and 2013) to project fertility by education levels from 2015 to 2030. Following the same procedure described in Lesotho, we fitted a model and Figure 4-7 shows the diagnostic plot of the baseline TFR over time (panel A) and the ratio of the other education categories to the baseline against the baseline TFR (panel B).

**Figure 4-7 Baseline TFRs and ratios to baseline TFRs, Namibia**




The results of the national TFRs are shown in Figure 4-8 together with the WPP. The percentage difference between the WPP and the SSP2 fitted points vary from 0.3 to 0.6 percent in 2015 and 2020, and 2.2 to 4.8 percent in 2025 and 2030 respectively. On the other hand, the SSP2-CER fitted points differ from the WPP estimates by 4 to 6 percent from 2015 to 2030. Therefore, there is little difference between our projected estimates and the WPP.

**Figure 4-8 Projected national TFR by education and WPP estimates in Namibia**




If education is to expand in a manner that the SSP2 scenario is met, then the WPP and the fitted TFRs using the SSP2 scenario will be little different from each other in future. The same applies to the case where the SSP2-CER scenario is satisfied. Both education scenarios are quite close to the WPP fertility estimates.

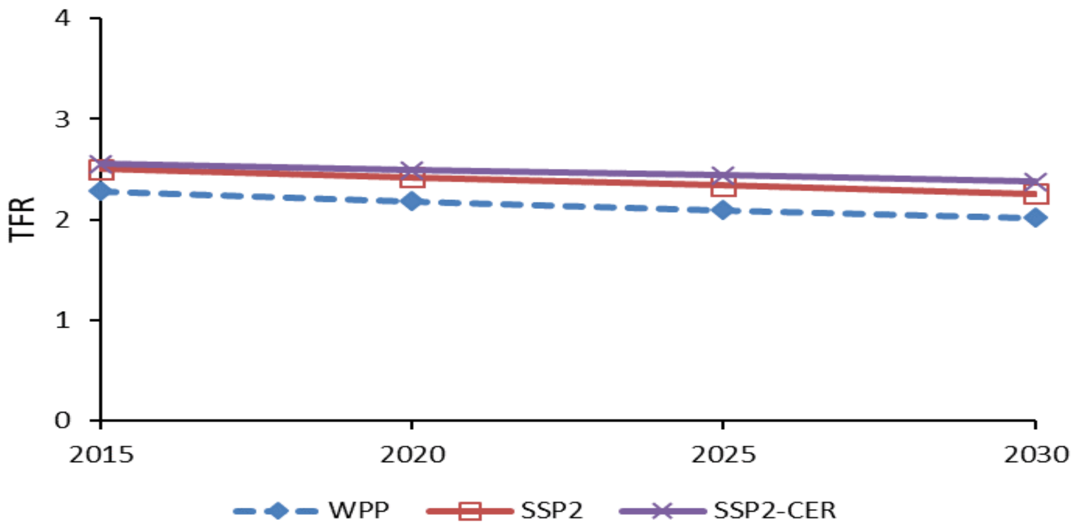
### 4.4.3 Fertility trends in South Africa

South Africa has only one DHS, which was carried out more than 15 years before 2015. One data point cannot be used to fit a model to predict future changes in fertility. We therefore had to seek some more information to fit this model. The first point of call was to choose a country which is at a later stage of its demographic transition. Colombia is a South American country which has sufficient data and is close to the end of its demographic transition, thus similar to South Africa which has one accessible DHS to date. We then assume that the slopes used to predict fertility change in Colombia are the same as those for South Africa, after which we fitted a model from these slopes and one data point for South Africa (Figure 4-9).

This approach does not seem to work since both the fitted values from both scenarios are at significantly different levels from the WPP estimates implying that the slopes for fertility decline in Colombia are less steep than the South African estimates.

It is misleading to use these countries to try and predict future changes in fertility for the country. It might be tempting to use the census data although in this study we need full birth history data as explained earlier, whereas censuses offer the summary birth information, therefore together with reasons stated earlier in Chapter 3, census data will not be used.

**Figure 4-9 Projected national TFR by education and WPP estimates, South Africa**



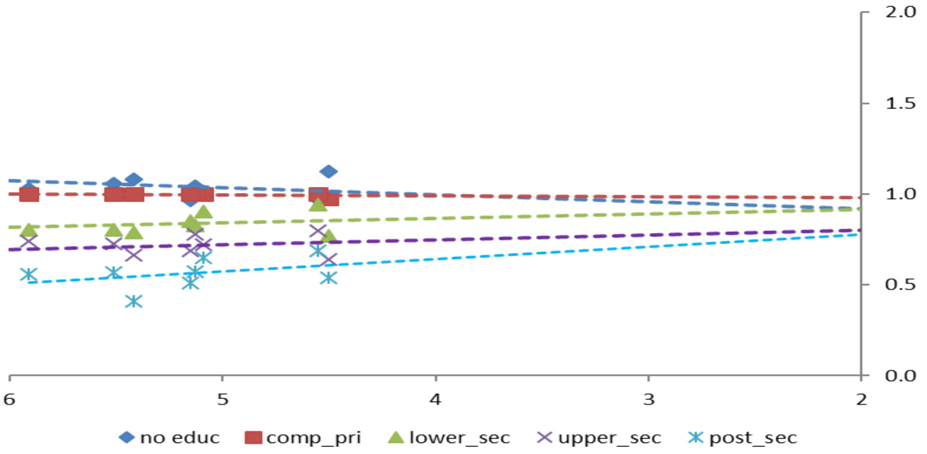

The results in Figure 4-9 indicate that each country has a unique pattern of fertility change because the two countries differ. The model, therefore, needs to be informed by the country's historic fertility data points. We cannot rely on other countries to effectively model fertility trends using this approach.

### 4.4.4 Fertility trends in Swaziland

Swaziland also has only one data point, and census data is not used because of the nature of the study and the information concerning detailed birth histories. In this case we combine all countries in the Southern African region, except South Africa, to fit a model that would predict fertility in the future for Swaziland on the assumption that it is a good representation of how fertility would change in the country. This exercise was performed to demonstrate how the model works when there is very limited data.

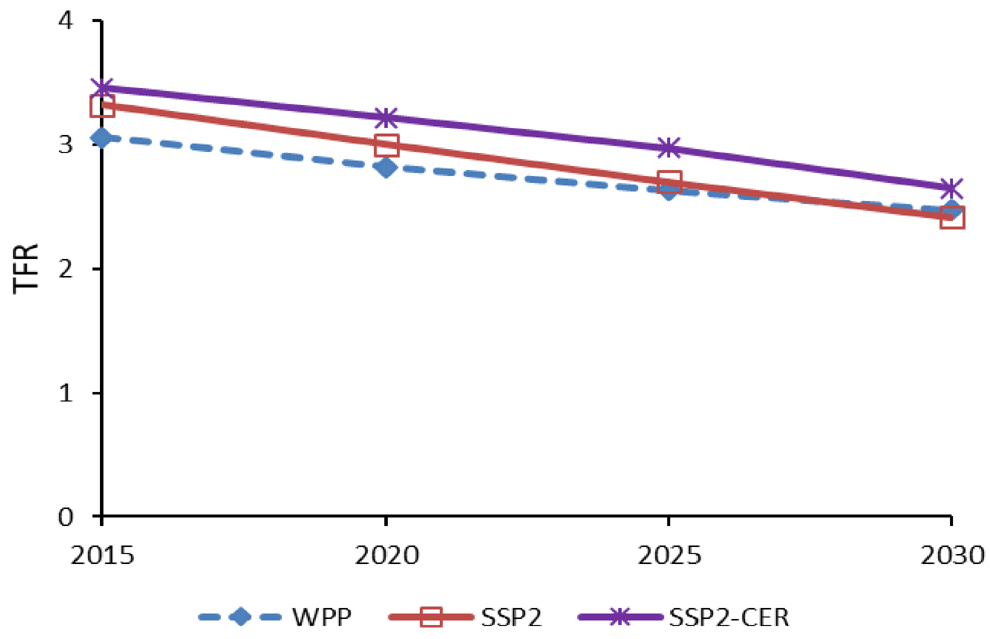
Although there were some points that were seen to be further from the straight line most the points approximate a straight line of the ratios to the baseline TFR (see Figure 4-10). Regardless of a weak linear relationship between the ratios to the baseline, the slopes derived from fitting a model using all the related countries are adapted to Swaziland.

**Figure 4-10 Baseline TFRs and ratios to baseline TFRs, Swaziland**




The results from projections using the two education scenarios are presented in Figure 4-11 together with the WPP fertility estimates. Major differences from the WPP are visible at the start of the projection period. The WPP and the fitted estimates also show different fertility decline rates. The fitted fertility rates decline faster than the WPP for both the scenarios of education attainment and the initial level of fertility in 2015 was seen to be much larger than the WPP fertility estimates. The levels and slopes of fertility decline differ from each other. Therefore, the model did not apply well in this case, and thus reliance on other country's data yields misleading results.

**Figure 4-11 Projected national TFR by education and WPP estimates, Swaziland**




### 4.4.5 Zimbabwe fertility trends

Zimbabwe had enough data points, and the model was fitted using these points and judgment in fitting the model to project the TFRs into the future. Figure 4-3 shows that fertility declined sharply from the 1988 to the 1994 ZDHS at all education levels, and stalled thereafter. A stalled pattern of fertility was observed in all education categories from the 1994 to the 2010-11 DHSs. Therefore, the inclusion of this DHS data point (1988 ZDHS) will greatly influence the slopes of fertility decline as they will tend to be steeper, and fertility will be declining faster when this point is included. Thus, the 1988 DHS was excluded when fitting a model to Zimbabwe.

Having decided on the DHS data points to include fitting the model to project fertility we then check if the assumption of linearity holds. Figure 4-12 shows a diagnostic plot to test whether this hold. There is a strong linear relationship between the ratios of each education category to the baseline (panel B), with all the points closer to the fitted lines. The baseline fertility (panel A) is changing in a manner that approximates a straight line as well. The main assumption for the model was satisfied when we applied the model to project fertility in Zimbabwe.

**Figure 4-12 Baseline TFRs against ratios of education levels, Zimbabwe**

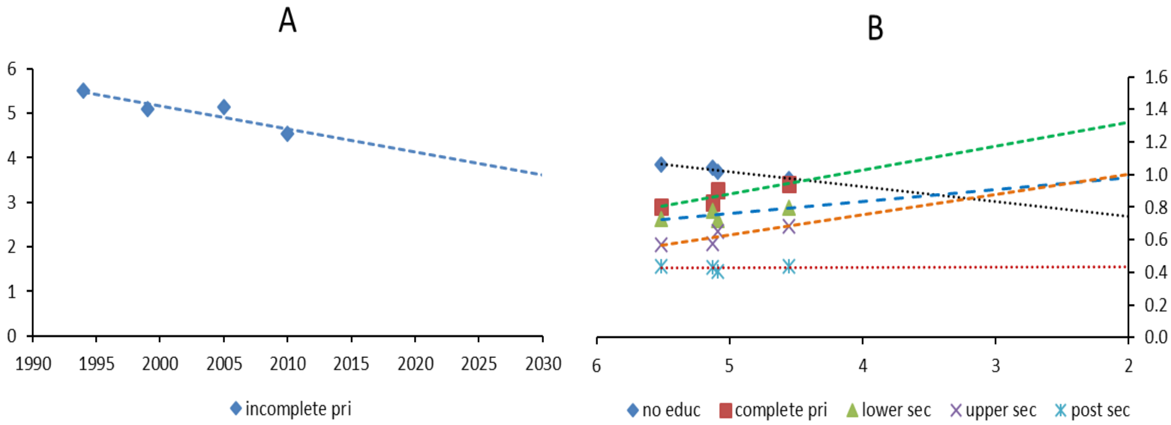
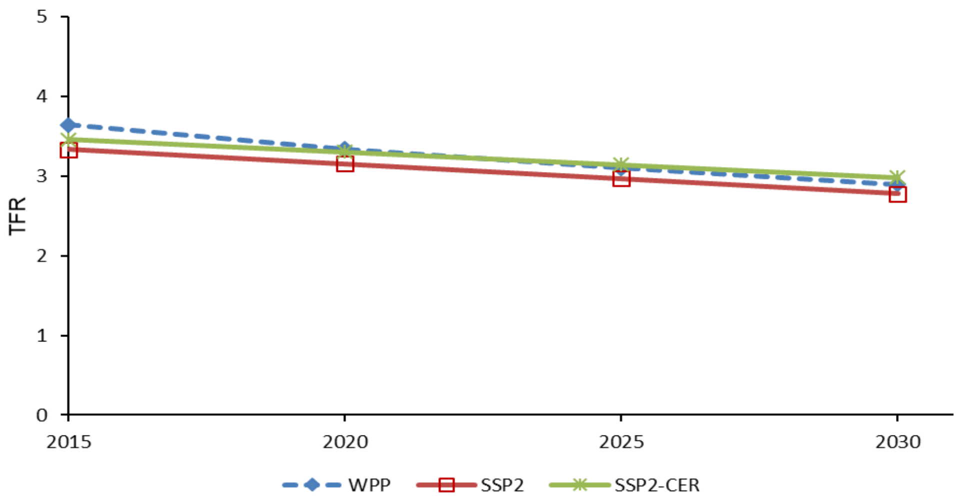



Figure 4-13 shows the results for TFRs projected from 2015 to 2030 in Zimbabwe. These show a downward trend in fertility in both cases for the projection period. The WPP TFRs are also displayed on the same graph (dotted line) and there is a consistent downward trend in fertility from 2015 to 2030.

The fertility estimates fitted using the SSP2-CER education scenario were very close to the WPP fertility estimates. The relative percentage difference between the two varied from a maximum of 3.7 percent in 2015 to a minimum of 1.2 percent in 2020 and finally by the year 2030 the relative percentage difference between the two projected estimates was 3 percent. The TFR estimates derived using the SSP2-CER education scenario are not different from the WPP TFR estimates for the whole projection period (2015-2030) considering small relative percentage differences (2.8 percent and less) for the whole period under consideration. The percentage difference between the SSP2 and the WPP fertility estimates varies from 8 percent in 2015 to 3 percent in 2030. Again, there is little difference between these projected and the WPP fertility estimates.

**Figure 4-13 Projected national TFR by education and WPP estimates, Zimbabwe**

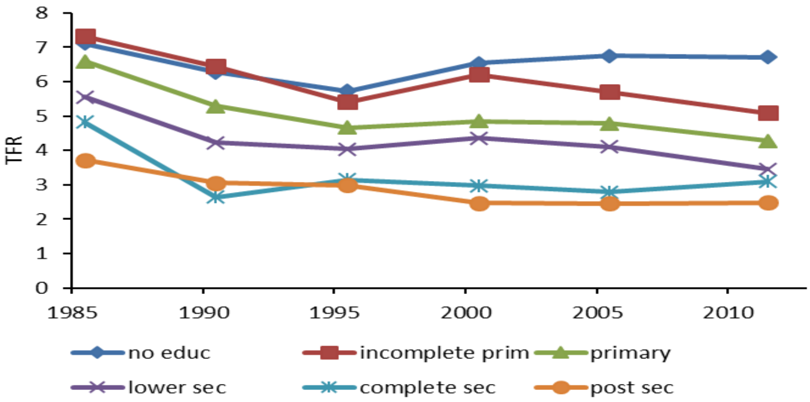


## 4.5 Model validation

The model was further validated using Kenya, a developing country in sub-Saharan Africa with six DHSs (1989, 1993, 1998, 2003, 2008 and 2014). The major reason why Kenya is chosen for this exercise, besides its similar fertility behaviour to the countries under study, is that it has more DHSs than are needed to fit the model. This exercise can be done while systematically omitting observed data points to investigate how well the model describes fertility trends when we vary the amount of data we have. This validation exercise addresses two questions: how well does the model fit using the first four DHSs compare with the last three observed fertility estimates? And does the model produce different results altogether when fewer DHS data points are used? Therefore, the work done here highlights how robust the proposed model is in projecting fertility.

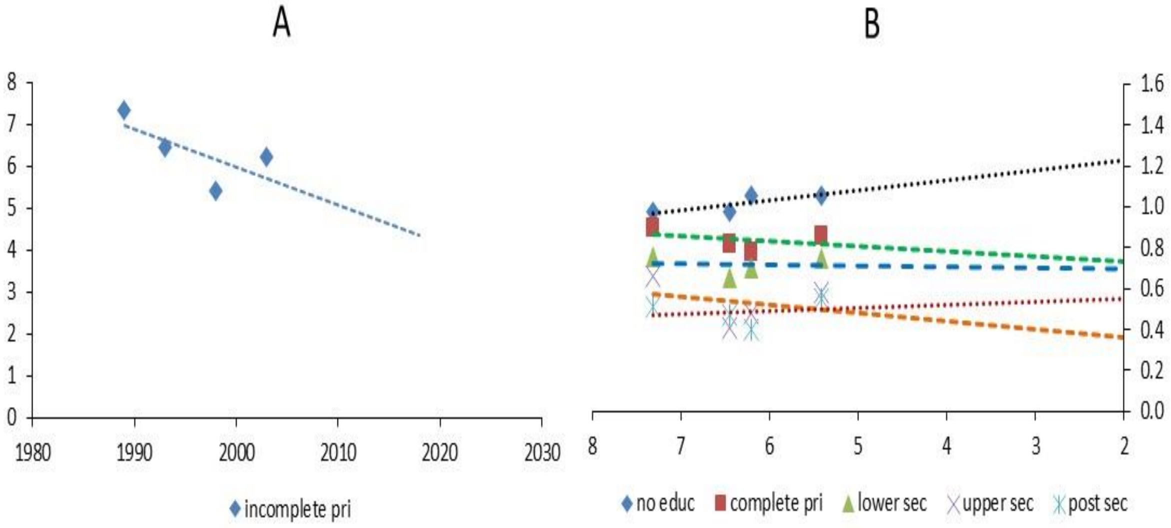
Fertility trends by education level in Kenya from the 1989 to 2014 DHS are shown in Figure 4-14. Fertility was higher among women with no education and the least educated, and lower among the more educated women (upper secondary and post-secondary). The fertility behaviour in Kenya is similar to that of countries under study, where fertility is negatively associated with educational attainment. The proposed model is then applied to Kenya using the first four DHSs to predict fertility in 2008 and 2014 (the last two DHSs). The predicted values are then compared to the observed values.

**Figure 4-14 Total fertility by education level, Kenya**




The left panel in Figure 4-15 shows the TFR for incomplete primary education (the baseline category). Fertility among women with incomplete primary was changing in a roughly linear fashion, despite the weak linear fertility estimates the method is explored. Therefore, we extrapolated the baseline TFRs further in time on the assumption that fertility will maintain a linear pattern. A diagnostic plot is shown in the right panel, namely, the ratios to the baseline TFRs were plotted against the baseline TFR, which also approximate linear relationships. This relationship is seen to be strongly linear in lower education categories and slightly weaker in higher education categories (upper- and post-secondary) due to reasons discussed in section 3.2. It is thus considered satisfactory to apply the model using these data points to predict fertility for the last two Kenyan DHSs.

**Figure 4-15 Baseline TFR over time and the diagnostic plot of ratios, Kenya**




The results from the fitted model using the first four Kenyan DHSs are presented in Table 4-2, where the absolute percentage difference between the fitted and the observed TFRs range from 4 percent to 6 percent in the 2008 and 2014 DHS respectively. These differences are small, and thus the fitted and the observed TFRs are close to each other. Evidently, the model produces consistent results with the observed TFRs, thus a good fit is obtained while using the four DHSs.

The exercise was further applied to fewer data points. In this case, we used the three DHSs after the first one to fit a model that is used to predict fertility for the two DHSs (2008 and 2014). The earliest DHS was excluded because it is further back in time and might not reflect future fertility trends. Fertility in the 1980s was still very high, and the inclusion of such data points will likely bias fertility downwards in the more distant future. Again, the results (Table 4-2) from the fitted model using the three DHS data points were consistent with the observed TFRs. The absolute percentage difference between the fitted and the observed value was approximately 4 percent in both DHS dates (2008 and 2014), which can be regarded as a better fit than the previous one (using 4 DHSs). This affirms the concept of excluding data that is further back in time in preference for the most recent data.

**Table 4-2 Observed and fitted TFRs for the last two DHSs, Kenya**

| *Year* | *Observed TFR* | *Fitted TFR with the first 4 DHSs* | *Fitted TFR with 3 DHS* |
|---|---|---|---|
| 2008 | 4.6 | 4.8 | 4.8 |
| 2014 | 4.4 | 4.7 | 4.6 |

In conclusion, the model was used to predict fertility in the last two DHSs in Kenya, which showed robustness as observed in the results. This exercise may be applied to many other countries with enough data points. Amongst other things, the method produces results which are similar to the WPP, which does not explicitly include educational dynamics as this approach does. Hence, the WPP may be seen as a simpler method, more parsimonious, but this model allows for the effects of educational distribution to be analysed.

# 5 DISCUSSION AND CONCLUSION

This chapter is divided into three sections. In section 5.1 we will discuss the results and their implications relative to the World Population Prospects (WPP), then present the limitations and problems or issues encountered in the study and method (section 5.2). We conclude in section 5.3 by discussing whether the objective was met and give recommendations for future lines of research in section 5.4.

## 5.1 Discussion of results

The first major task was to investigate how well the historical educational attainment distributions from the IIASA and DHS match one another. Historical estimates from these two were similar, which indicates that both these two data sources describe the same populations. These distributions are in turn used to derive national fertility rates from the DHS age-education specific fertility rates compared to the observed fertility rates. The resulting national fertility rates were similar as well and not very different from the observed TFRs. Thus, the IIASA and the DHS education attainment distributions therefore describe the Southern African populations in a similar way to a larger extent as suggested by the results, since the difference between the two was insignificant (mostly less than 5percent).

The results also show that fertility change is broadly inversely related to educational attainment of women. This is supported by the findings that as fertility declined in the region, the proportion of women in higher education increased consistently. Fertility tends to be higher among less educated women relative to more educated women. There is evidence that education composition change has a negative effect on fertility change, as shown when we fixed fertility in all education categories and allowed the education composition to change over time. The resulting national fertility rates highlighted a decline in fertility over the projection period. These results are consistent across Southern Africa. Overall, fertility decline is not only due to a change in the education component of women but also due to the change in fertility levels among women’s categories of education.

The model for projecting fertility by education categories is best fitted using the country-specific DHS data points. For countries, which have only one DHS data point (South Africa and Swaziland) the model resulted in inconsistent fertility patterns. However, for other countries (Lesotho, Namibia and Zimbabwe) which have sufficient

DHS data points, the model produces results consistent with the WPP fertility estimates. A key assumption made without justification is that the trend in the ratio of TFRs by education level to those with "incomplete-primary" is linear. While this might be a reasonable assumption for short-term projections (particularly if one has past data for a similar number of years as the projection period) in general it is a problematic assumption for long-term projections in that there is nothing to prevent implausibly low (or high) projected TFRs. Thus in Figure 4-5 which shows ratios for Lesotho, we see the projected ratio for those with no education reaching that of those with post-secondary education (implying a TFR of 0.84) while at the same point the TFR of the post-secondary education is more than twice as high.

The validation exercise using the Kenyan DHSs affirmed the model's robustness, and may be applied to many other countries since the predicted and the observed fertility rates were very similar. Alterations were made by using fewer data points when fitting a model, which still yielded plausible projections, which cements the notion that at least three DHSs yield plausible results if the baseline TFRs are strongly linearly related to the ratios of TFRs to the baseline education category. We observed a weaker linear trend when four data points were used relative to three. That explains why the 2014 fitted estimate using three DHS was closer to the observed value as opposed to using four. Judgment must be used as well in fitting a model as to which points are relevant to include. This was done when fitting a model to countries with more data where we excluded the older DHSs to avoid misleading results due to rapid fertility decline which existed in the past.

Fertility has since been declining slowly or stalled in many countries of Sub-Saharan Africa since the late 1990s. The stalling fertility trends were seen in the later DHSs for Namibia and Zimbabwe. The model was adjusted for both countries by excluding the more distant surveys. This adjustment is, however, subject to availability of data to comprehend if fertility had stalled in a country.

The United Nations used a first-order autoregressive model but only for countries that have reached replacement level fertility in projecting fertility, and their model predicted that fertility will be oscillating around replacement-level fertility (2.1 children per woman). However, for countries that are still undergoing a fertility transition, a Bayesian Hierarchical Model (BMH) was used (see section 2.3.1). There is no evidence that their model addresses the issue of stalling fertility trends and so is the proposed model, meaning that although the model seemed to be working well there is a lot to be done so that the issue of stalling fertility trends in Southern Africa is factored in. The results predicted by the proposed projection technique are plausible for the projection interval and may be misleading when applied in long-term projections. This new technique allows for education to be analysed, thus may be useful in this sense.

## 5.2 Limitations

In this study, only a few countries were considered because of limited time. There is still more to be done by other researchers, applying the model to many other countries, to check if it gives plausible results as well. The model needs at least three DHS data points for a specified country to be robust. However, if there are two data points the method can still be used, but the diagnostic plots will be meaningless to inspect and interpret because basically two points will always lie in a straight line.

A major limitation to the study is the availability of DHS data in countries of inquiry. This is the only source data used because we require the full birth-history data to be able to compute fertility rates by tabulating women and exposure by age, while other sources such as censuses were reviewed but mostly offered summary information which was less useful. We restricted the study to only one source of data partly to avoid differentials that might arise due to different sources. The other reason is that it was simpler to use the *tfr2* STATA macro-command to compute TFRs using the standard recode variables from DHS as compared to the censuses or other surveys as one would have to recreate new variables to allow for such tools to be used. This process may give rise to other sources of error in our estimates. South Africa and Swaziland both have only one DHS publicly available to date. In South Africa, there are two DHSs to date, although access to the 2003 DHS is restricted because of the inconsistencies in the data exhibited. Thus, we only relied on one SADHS.

In South Africa and Swaziland, we assumed the fertility trends to fit the model, this evidently became problematic and the model did not fit well. Furthermore, even if there are sufficient DHSs for other countries, fragmentation of data into different education categories results in smaller sample sizes as well. Hopefully in future studies, larger samples will be attainable. However, such a solution is governed by financial and other resource constraints.

If there were enough DHSs for Swaziland and South Africa, the model is expected to work better and give a good indication of how fertility is likely to change. In addition, this model was also applied to other countries with more DHSs and were not part of the study (Colombia and Kenya). It produced plausible results when validated.

However, it can also be applied to many more countries to be surer of how the model applies in other countries. Should it prove robust, this model may be a preferred alternative of projecting fertility by educational attainment, although there are some constraints to this model.

One of the major constraints on our method is how well the educational composition is modelled by IIASA data, as even if women are split by education levels in surveys, this only captures a snapshot of what is happening at one point in time. Younger women tend to change education levels quite rapidly. These distributions might not represent the populations' educational composition well, thus we are limited to the information we have at hand.

Reconstructing fertility trends using the DHS data, which is cross-sectional, we implicitly assume that the population characteristics remain fixed in that period. From the findings, such as population characteristics, for example, the educational composition continued to change over time, thus we are limited to such assumptions to calculate fertility rates using the DHS data. Women's educational attainment proportions are fixed at the date of the survey. However, these are subject to change rapidly for younger women (15-19 years) as they can progress to higher education levels, unlike those over 35 years. For example, we would not know at what stage women would be a year later, which tends to bias estimates derived for such age-groups. There is nothing we can do about that since the DHS only shows a snapshot of the population at that point in time.

The method relies on many assumptions compared to the Bayesian Hierarchical Model (BHM), which is currently used in the World Population Prospects (WPP), and which might be a major weakness to such a modelling technique. However, there are advantages of using such a projection model: firstly, the approach allows policy changes in the field of education to be allowed for; and secondly, it allows us to test further the differential effects of those policy changes on fertility. In addition, it has fewer parameters, which must be estimated to fit a projection model relative to the BHM. Only two parameters are estimated from the DHS data in our model, which are the slopes and intercepts of fertility patterns for each country. In contrast, the BHM is more complex since it uses a probabilistic model to estimate six parameters and by including the country-specific and global experience make the process even more difficult to understand and replicate. Probabilistic projections (for example, the BHM) have an advantage over this model in the sense that they allow projections for extended periods (up to 85 years), and allow for uncertainty bounds to calculated. This aids in estimating the uncertainty that exists in such long and ultra-long projections.

## 5.3 Conclusion

In the study, we set out to project fertility by including education composition change across five Southern African countries for a period from 2015 to 2030 using the DHS data. Although, the DHS was chosen as the ultimate data source because it offers detailed birth history information, it has its flaws. These ranged from the time it takes in conducting surveys to release of data and population representativeness. However, plausible age and fertility structures were evident in Southern Africa. The data evaluation exercise was useful in the sense that it helps explain some of the results we get from such data. We also extracted the education proportions from IIASA and to ensure comparability, the DHS data was harmonised and matched the six education categories from IIASA. To a greater extent, this exercise proved to be a success since the age-education structure of the two sources is very similar.

In handling the DHS data for analysis, a very useful STATA macro-command (*tfr2*) made it easier to calculate fertility rates and the process is easily replicable, as this was done several times and still yielded consistent TFR results. Patterns of fertility change by education levels were determined, as were the national fertility estimates. The reconstructed fertility rates validated the central importance of education composition change on fertility. This finding has significant implications for the study of fertility trends by education level.

Fertility decline in Southern African countries is strongly associated with education composition change: when the proportion of women in higher education level increases over time, fertility declines consistently. However, education composition change is not sufficient to bring about overall fertility decline.

A projection model was proposed in this dissertation, which concept seems to work well under certain conditions. In countries with more than one DHS data point (Lesotho, Namibia and Zimbabwe) the concept produces reliable fertility estimates over a 15-year projection period (2015-2030), which is judged with what is known at present. However, the model fails for countries with insufficient data points (South Africa and Swaziland). The assumption that the ratios of TFRs for each education level are linearly related to the baseline TFRs must be satisfied for this method to produce reasonable results, and when it fails the method is inapplicable. The approach is robust when more data points are present and an approximation of a straight line of the ratios to the baseline TFRs produce consistent projected fertility rates. This assumption of linearity is problematic when we consider long projection intervals, thus at this point the method can only be applicable for short period projections. The incomplete-primary education category was characterised by small numbers of women in Southern Africa, this is associated with implausible ratios in this education category that were observed relative to the baseline TFRs.

The validation exercise performed and presented in section 4.5 shows that the projected fertility estimates in Kenya are not very different from the observed fertility rates of the 2008 and 2014 DHSs. Therefore, by fitting the projection model we could reproduce the DHS fertility rates and this further validated the projection technique developed in this study. This also supports the assertion that the model is better fitted using sufficient country-specific data provided the assumptions are not violated.

We learned that in Southern Africa, women's education plays an important role on fertility decline. In addition, there is pronounced heterogeneity that exists between different education categories that needs to be considered when projecting fertility. The findings that different education policies lead to different fertility trends are affirmed by the study. Modelling of fertility trends using a new approach that was validated gives evidence of how well the model predicts fertility for countries. This model is developed from an understanding of how fertility changes in populations in one education category relative to others. This brings us to the conclusion that it is not possible to replicate the WPP fertility projections since they are discrepancies between the two, but the study resulted in consistent fertility estimates to the WPP. Both projection methods rely on assumptions but the method presented here explicitly factored in the effects of women's education attainment in fertility dynamics, and in future education must be considered in fertility projections to eliminate some of the differentials that may arise due to education composition change.

## 5.4 Recommendations

This study has many lessons to take note of for further work on fertility studies and, while acknowledging the limitations of the study, there are recommendations for future studies in this field. The DHS data were the only source of data used for analyses, although not accessible for all countries in Southern Africa. In future, other sources such as census data might be included in a model, further adapted to incorporate this data. The exclusion of countries with insufficient data (South Africa and Swaziland) would make the modelling process much easier, but they were included in our study as they point to the consequences of inadequate data. This may also influence other researchers on how to best apply the model using other sources of data and may give rise to the development and application of demographic estimation tools where data is inaccessible.

The *tfr2* STATA macro-command was developed and useful in computation of measures of fertility in DHS data, where the dates are coded in century month code (CMC). We recommend that this tool be expanded to be compatible with data from other sources so that it may be easy to explore the method into greater heights. Data sources in Southern Africa still needs improvement, and therefore countries must invest in improving data qualities.

We only focused on incorporating education attainment in projecting fertility, therefore, we recommend the model to be implemented using other sources of heterogeneity such as type of place of residence and religion. The other major issue that we recommend being further incorporated is the issue of stalling fertility trends in projection models. In our study this was not controlled for but by excluding the more distant surveys the model yielded plausible results. However, a gradual decrease in fertility is inevitable in the long-run. Therefore, further studies may incorporate these stalls in fertility trends as this might raise new insights about the fertility trajectory in sub-Saharan Africa.

More methodological work on how robust the projection model is needed, including further developing the model and adding some probabilistic features that will allow uncertainty bounds to be estimated and in-depth exploration on how well the model predicts fertility is required. These will ensure that it is applicable even in long or ultra-long projections.

It is of great significance that more studies of this nature be done to inform governments and organisations on how to control fertility levels by investing in women's education. Different education scenarios are presented and each inform on how fertility may change if one education policy was to be adapted. The question as to whether the model improves the accuracy of projections will be up to other researchers to decide, but the model allows for education which is a significant driver of fertility change to be analysed.

## REFERENCES


Akmam, W. 2001. "Maternal education as a strategy for children's survival and health in developing countries, with special reference to Bangladesh". *Eubios Journal of Asian and International Bioethics*, **11**(2001):76-78.

Akmam, W. 2002. "Women's education and fertility rates in developing countries, with special reference to Bangladesh". *Eubios Journal of Asian and International Bioethics*, **12**(4):138-143.

Al Riyami, A., Afifi, M. and Mabry, R.M. 2004. "Women's autonomy, education and employment in Oman and their influence on contraceptive use". *Reproductive health matters*, **12**(23):144-154.

Alkema, L., Raftery, A.E., Gerland, P., Clark, S.J., Pelletier, F., Buettner, T. and Heilig, G.K. 2011. "Probabilistic Projections of the Total Fertility Rate for All Countries". *Demography*, **48**(3):815-839.

Asiimwe, J.B., Ndugga, P. and Mushomi, J. 2013. Socio-demographic factors associated with contraceptive use among young women in comparison with older women in Uganda. Calverton, Maryland, USA: ICF International. http://dhsprogram.com/pubs/pdf/WP95/WP95.pdf.

Bauer, R., Potancokova, M., Goujon, A. and Kc, S. 2012. Populations for 171 Countries by Age, Sex, and Level of Education around 2010: Harmonized Estimates of the Baseline Data for the Wittgenstein Centre Projections. Laxenburg, Austria: Wittgeinstein Centre. http://pure.iiasa.ac.at/10259/1/IR-12-016.pdf.

Blacker, J. 1994. "Some thoughts on the evidence of fertility decline in Eastern and Southern Africa". *Journal of the American Statistical Association* **20**(1):200-205.

Bongaarts, J. 1978. "A Framework for Analyzing the Proximate Determinants of Fertility". *Population and Development Review*, **4**(1):105-132.

Bongaarts, J. 2006. "The Causes of Stalling Fertility Transitions". *Studies in Family Planning*, **37**(1):1-16.

Bongaarts, J. 2010. "The causes of educational differences in fertility in Sub-Saharan Africa". *Vienna Yearbook of Population Research*, **8**(2010):31-50.

Bongaarts, J. and Feeney, G. 1998. "On the Quantum and Tempo of Fertility". *Population and Development Review*, **24**(2):271-291.

Bongaarts, J., Frank, O. and Lesthaeghe, R. 1984. "The Proximate Determinants of Fertility in Sub-Saharan Africa". *Population and Development Review*, **10**(3):511-537.

Bongaarts, J. and Potter, R. 1985. "Fertility, Biology, and Behavior. An Analysis of the Proximate Determinants". *Journal of the American Statistical Association*, **57**(2):304-306.

Cai, F. 2010. "Demographic transition, demographic dividend, and Lewis turning point in China". *China Economic Journal*, **3**(2):107-119.

Caldwell, J.C. and Caldwell, P. 1987. "The Cultural Context of High Fertility in sub-Saharan Africa". *Population and Development Review*, **13**(3):409-437.

Caldwell, J.C., Orubuloye, I.O. and Caldwell, P. 1992. "Fertility Decline in Africa: A New Type of Transition?". *Journal of the American Statistical Association*, **18**(2):211-242.

Chang, J.-S. 1996. "What Do Education and Work Mean? Education, Nonfamilial Work/Living Experiences and Premarital Sex for Women in Taiwan". *Journal of Comparative Family Studies*, **27**(1):13-40.

Chaudhury, R.H. 1978. "Female status and fertility behaviour in a metropolitan urban area of Bangladesh". *Population Studies*, **32**(2):261-273.

Chesnais, J.-C. 1992. *The Demographic Transition: Stages, Patterns, and Economic Implications,* Oxford: Oxford University Press Catalogue.

Cleland, J. 2002. "Education and future fertility trends, with special reference to mid-transitional countries". *Completing the fertility transition*, **1**(2002):183-207.

Cleland, J. and Jejeebhoy, S. 1996. "Maternal schooling and fertility: evidence from censuses and surveys". *In:* ROGER JEFFERY and ALAKA, M BASU (eds.) *Girls' schooling, women's autonomy and fertility change in South Asia.* New Delhi, India: Sage Publications, pp. 72-106.

Cochrane, S.H., Khan, M.A. and Osheba, I.K.T. 1990. "Education, income, and desired fertility in Egypt: A revised perspective". *Economic Development and Cultural Change*, **38**(2):313-339.

Cochrane, S.H., Pandey, G., Pandey, J., Rice, F., Lanctot, C., Garcia-Devesa, C., Khan, M., Osheba, I., Robinson, W. and Cleland, J. 1979. "Fertility and education: What do we really know?". *Pop Cen News Letter*, **5**(4):1-5.

Davis, K. and Blake, J. 1956. "Social Structure and Fertility: An Analytic Framework". *Economic Development and Cultural Change*, **4**(3):211-235.

Dyson, T. 2011. "The Role of the Demographic Transition in the Process of Urbanization". *Population and Development Review*, **37**(2011):34-54.

Easterlin, R.A. 1983. "Modernization and fertility: a critical essay". *In:* BULATAO RA and LEE, RD (eds.) *Determinants of fertility in developing countries: Fertility regulation and institutional influences.* New York: Academic Press, pp. 562-86.

Hill, K. and Dorrington, R.E. 2013. "Introduction to migration analysis". *In:* IN MOULTRIE TA RE DORRINGTON, AG HILL, K HILL, IM TIMÆUS AND B ZABA (ed.) *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population, pp.

Hindin, M.J. 2000. "Women's autonomy, women's status and fertility-related behavior in Zimbabwe". *Population Research and Policy Review*, **19**(3):255-282.

Ibisomi, L., Williams, J., Collinson, M.A. and Tollman, S. 2014. "The stall in fertility decline in rural, northeast, South Africa: the contribution of a self-settled, Mozambican, refugee sub-population". *Etude de la Population Africaine*, **28**(1):590.

Jejeebhoy, S.J. 1995. *Women's education, autonomy, and reproductive behaviour: Experience from developing countries* [Online]. Available: https://ideas.repec.org/b/oxp/obooks/9780198290339.html [Accessed 16 June 2016].

Johnson, K., Abderrahim, N. and Rutstein, S.O. 2011. Changes in the direct and indirect determinants of fertility in sub-Saharan Africa. Calverton, Maryland, USA: ICF Macro. http://dhsprogram.com/pubs/pdf/AS23/AS23.pdf.

Jolly, C.L. and Gribble, J.N. 1993. "The proximate determinants of fertility". *Demographic change in sub-Saharan Africa*, **1**(26):68-116.

K.C., S. and Abel, G.J. 2015. *Wittgenstein Centre Data Explorer Version 1.2* [Online]. Available: http://www.wittgensteincentre.org/dataexplorer [Accessed 27 February 2016 2016].

Kc, S. and Lutz, W. 2014. "The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100". *Global Environmental Change*, **42**(2014):181-192.

Kirk, D. 1996. "Demographic Transition Theory". *Population Studies*, **50**(3):361-387.

Lee, R.D. and Tuljapurkar, S. 1994. "Stochastic Population Forecasts for the United States: Beyond High, Medium, and Low". *Journal of the American Statistical Association*, **89**(428):1175-1189.

Lesthaeghe, R., Ohadike, P.O., Kocher, J. and Page, H.J. 1981. "Child spacing and fertility in sub-Saharan Africa: an overview of issues". *Studies in Populatio*, **1**(1981):3-23.

Lutz, W. and Goujon, A. 2001. "The World's Changing Human Capital Stock: Multi-State Population Projections by Educational Attainment". *Population and development review*, **27**(2):323-339.

Lutz, W. and Goujon, A. 2003. "The World's Changing Human Capital Stock: Multi-State Population Projections by Educational Attainment". *Vienna Yearbook of Population Research*, **1**(2):19-33.

Lutz, W. and Kc, S. 2013. "Demography and Human Development: Education and Population Projections". *International Institute for Applied Systems Analysis*, **1**(4):1-18.

Martin, T.C. 1995. "Women's Education and Fertility: Results from 26 Demographic and Health Surveys". *Studies in Family Planning*, **26**(4):187-202.

Miller, T. 2006. *Demographic models for projections of social sector demand*: United Nations Publications.

Ministry of Health [Lesotho] and Icf International. 2016. Lesotho Demographic and Health Survey 2014. Maseru, Lesotho: Ministry of Health and ICF International.

Moultrie, T.A. 2013. " Introduction to fertility analysis". *In:* IN MOULTRIE TA RE DORRINGTON, AG HILL, K HILL, IM TIMÆUS AND B ZABA (ed.) *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population, pp.

Moultrie, T.A., Sayi, T.S. and Timæus, I.M. 2012. "Birth intervals, postponement, and fertility decline in Africa: A new type of transition?". *Population Studies*, **66**(3):241-258.

Raftery, A.E., Alkema, L. and Gerland, P. 2014. "Bayesian Population Projections for the United Nations". *International Institute for Applied Systems Analysis*, **29**(1):58-68.

Rogers, R.G. and Hackenberg, R. 1987. "Extending epidemiologic transition theory: a new stage". *Social biology*, **34**(3-4):234-243.

Rustein, S.O. and Rojas, G. 2006. Guide to DHS Statistics. Calverton, Maryland: Demegraphic and Health Surveys: ORC Macro. https://www.google.com/url?q=http://www.dhsprogram.com/pubs/pdf/DHSG1/Guide_to_DHS_Statistics_29Oct2012_DHSG1.pdf.

Sambisa, W. 1996. Contraceptive use dynamics in Zimbabwe: Discontinuation, switching, and failure. Calverton, Maryland, USA: Macro International. http://dhsprogram.com/pubs/pdf/FA15/FA15.pdf.

Schoumaker, B. 2013. "A Stata module for computing fertility rates and the TFRs from birth histories: tfr2". *Demographic Research* **28**(1):1094-1135.

Schoumaker, B. 2014. Quality and consistency of DHS fertility estimates, 1990 to 2012. Rockville, Maryland, USA: ICF International. http://dhsprogram.com/pubs/pdf/MR12/MR12.pdf.

Shapiro, D. 2017. "Linkages between Education and Fertility in Sub-Saharan Africa". *AFD Research Paper Series*, **56**(2017).

Shapiro, D. and Gebreselassie, T. 2013. "Fertility transition in sub-Saharan Africa: falling and stalling". *African Population Studies*, **23**(1).

Speringer, M., Goujon, A., Kc, S., Eder, J., Potancokova, M. and Bauer, R. 2015. Validation of the Wittgenstein Centre Back-projections for Populations by Age, Sex, and Six Levels of Education from 2010 to 1970. Laxenburg, Austria: International Institute for Applied Systems Analysis. http://pure.iiasa.ac.at/11680/1/IR-15-008.pdf.

Swartz, L. 2009. "Fertility transition in South Africa and its implications on the four major population groups". *Completing the Fertility Transition*, **1**(48):487-500.

Teitelbaum, M.S. 1975. "Relevance of Demographic Transition Theory for Developing Countries". *Science*, **188**(4187):420-425.

The Demographic and Health Surveys Program. 2016. *Datasets* [Online]. Available: https://dhsprogram.com/data [Accessed 20 January 2016].

Timæus, I.M., Dorrington, R.E. and Hill, K.H. 2013. "Introduction to adult mortality analysis". *In:* MOULTRIE TA RE DORRINGTON, AG HILL, K HILL, IM TIMÆUS AND B ZABA (ed.) *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population, pp.

United Nations Department of Economic and Social Affairs Population Division. 2015. *World Population Prospects: The 2015 Revision, custom data acquired via website* [Online]. Available: http://esa.un.org/unpd/wpp/DataQuery/ [Accessed 21 February 2016.

United Nations Department of Economic and Social Affairs: Population Division. 2015. World Population Prospects: The 2015 Revision, Methodology of the United Nations Estimates and Projections. New York: United Nations.

Westoff, C.F. and Cross, A.R. 2006. The stall in the fertility transition in Kenya. Calverton, Maryland, USA: ORC Macro. http://dhsprogram.com/pubs/pdf/AS9/AS9.pdf.

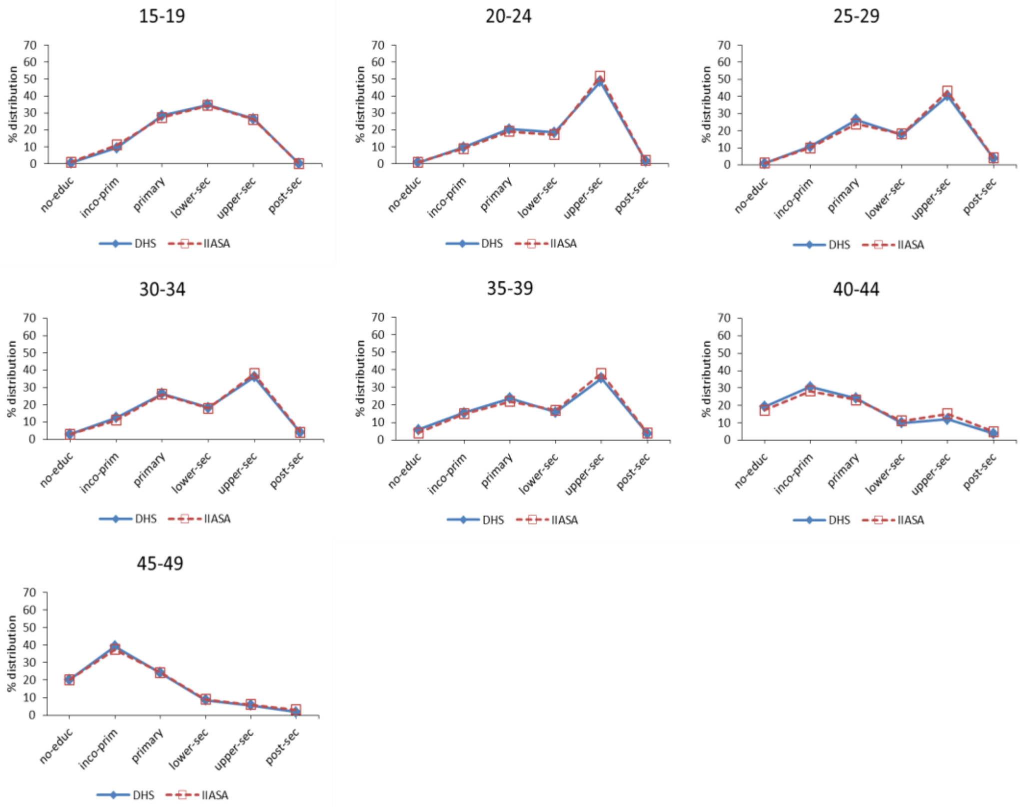
Wittgenstein Centre for Demography and Global Human Capital. 2015. *Wittgenstein Centre Data Explorer Version 1.2.* [Online]. Available: http://www.wittgensteincentre.org/dataexplorer [Accessed 16 February 2016].

Wittgenstein Centre for Demography and Global Human Capital. 2016. *Scenario Definitions* [Online]. Available: http://www.oeaw.ac.at/vid/dataexplorer/ [Accessed 6 June 2016].

# APPENDICES

## Appendix A Comparison of the education percentage distributions (DHS vs. IIASA)

### Figure 5-1 Percentage education distribution, Zimbabwe 2005-06 DHS




### Figure 5-2 Percentage education distribution, Zimbabwe 1999 DHS

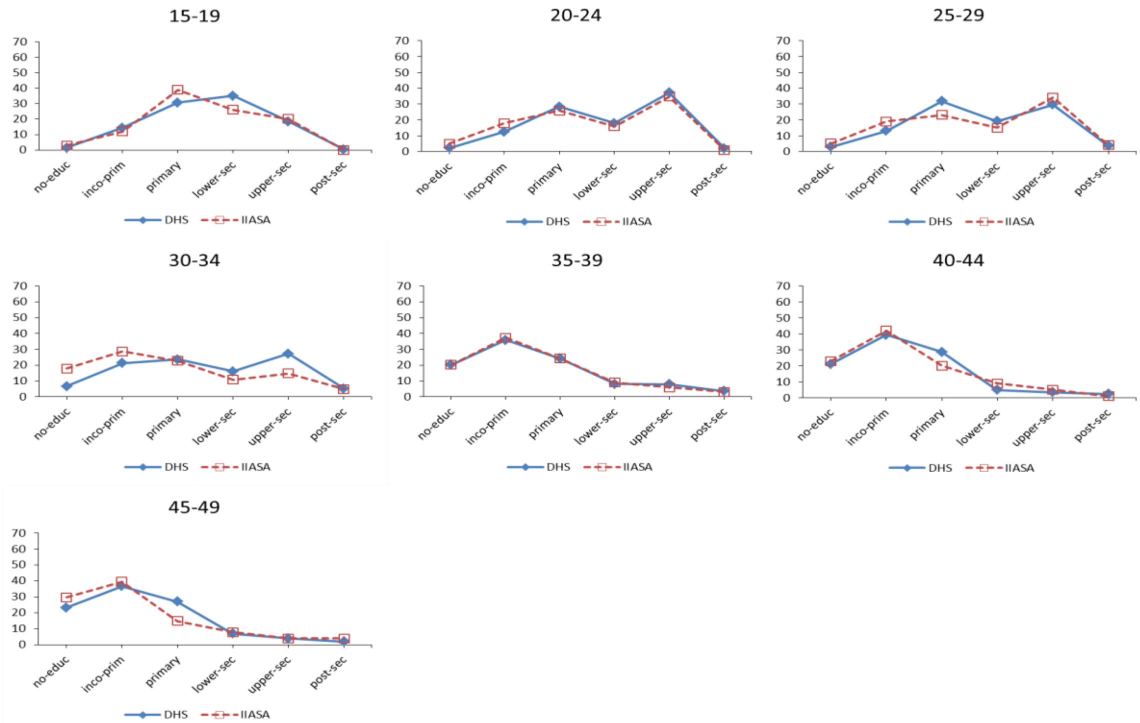


**Figure 5-3 Percentage education distribution, Zimbabwe 1994 DHS**

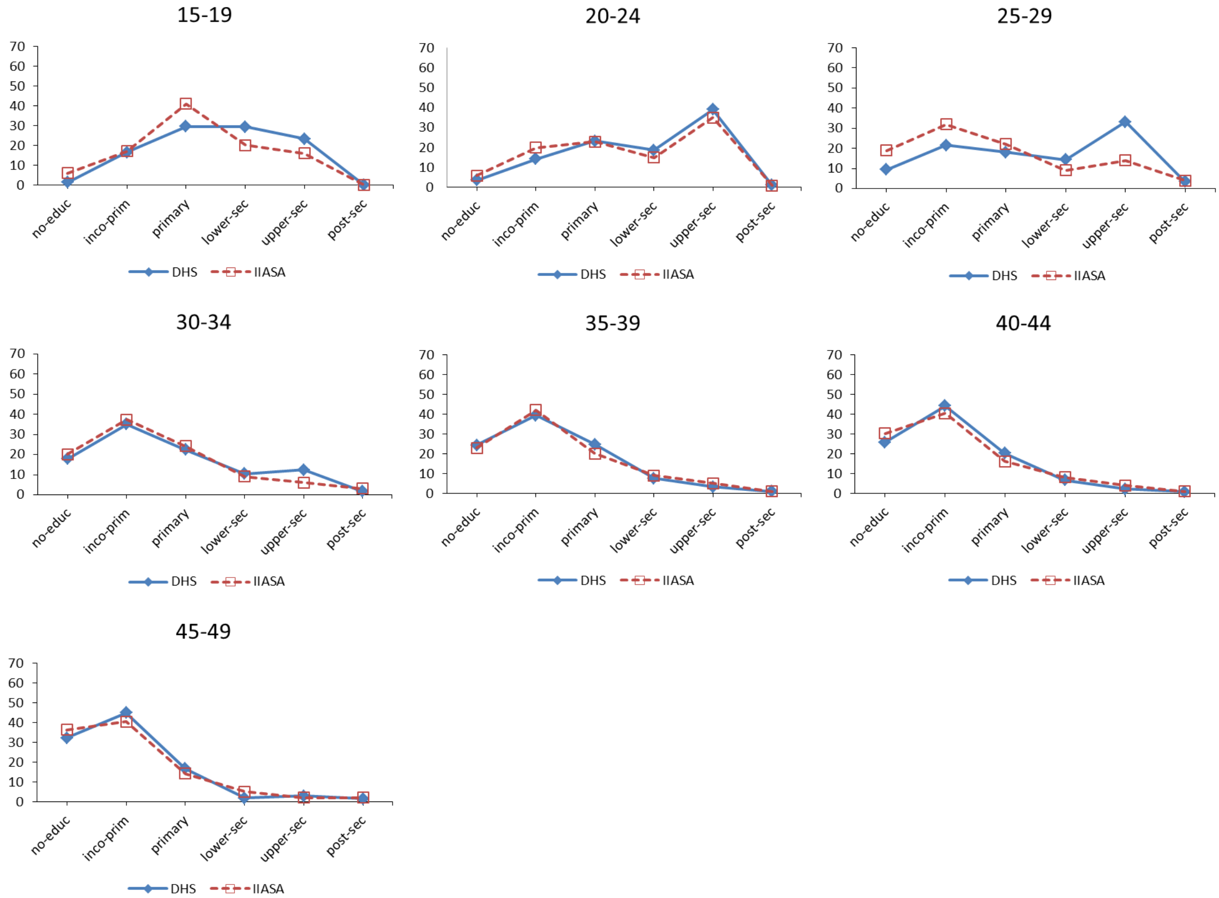


Figure 5-4  Percentage education distribution, Zimbabwe 1988 DHS

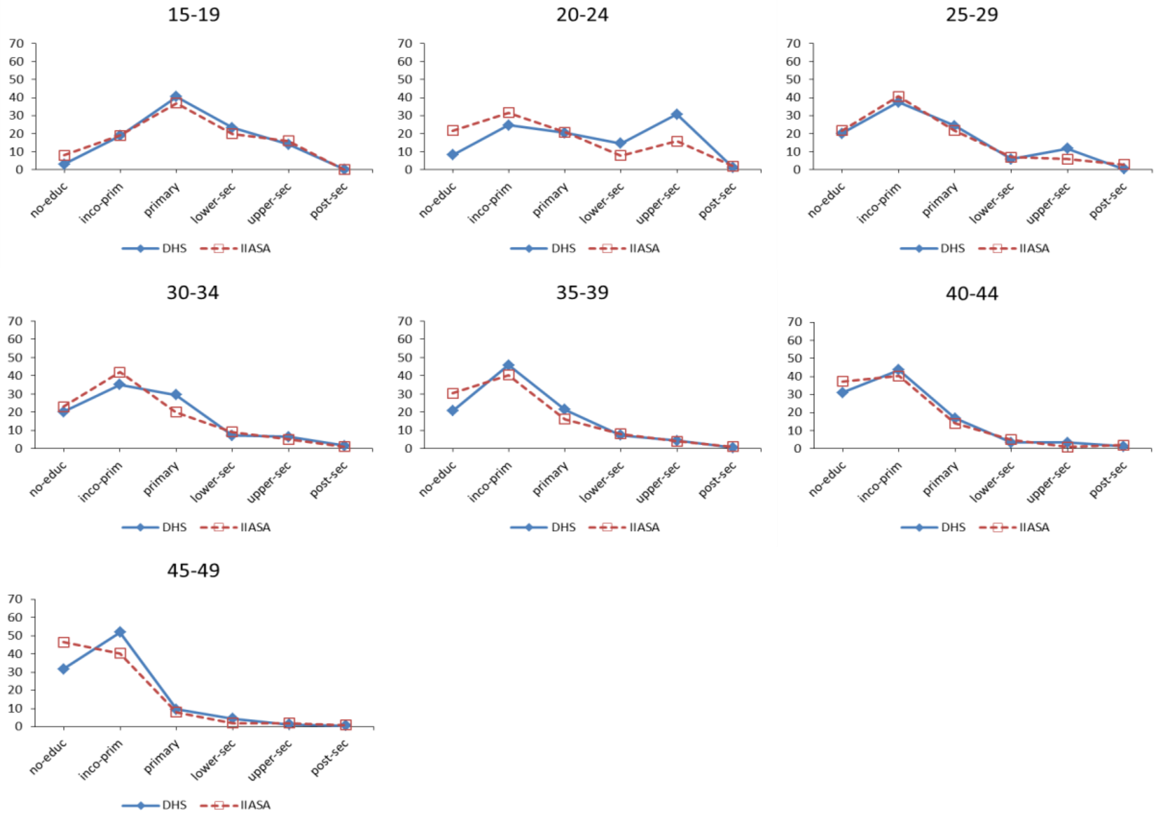
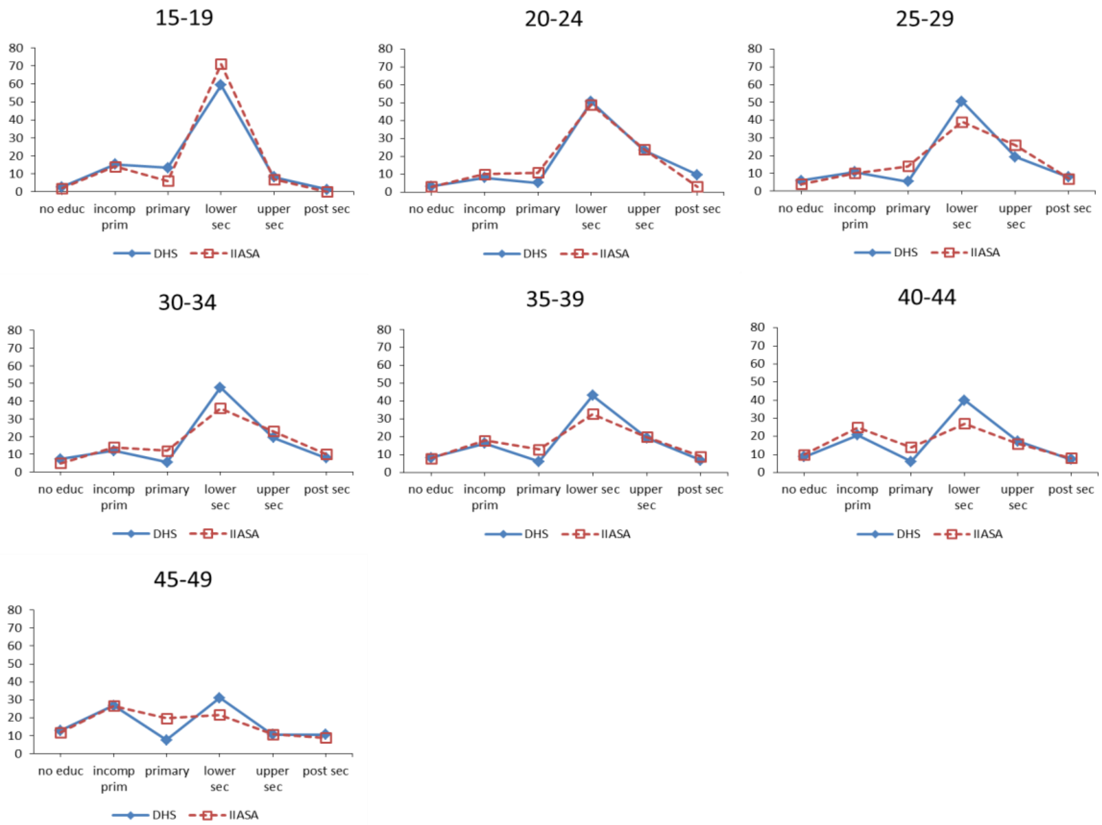
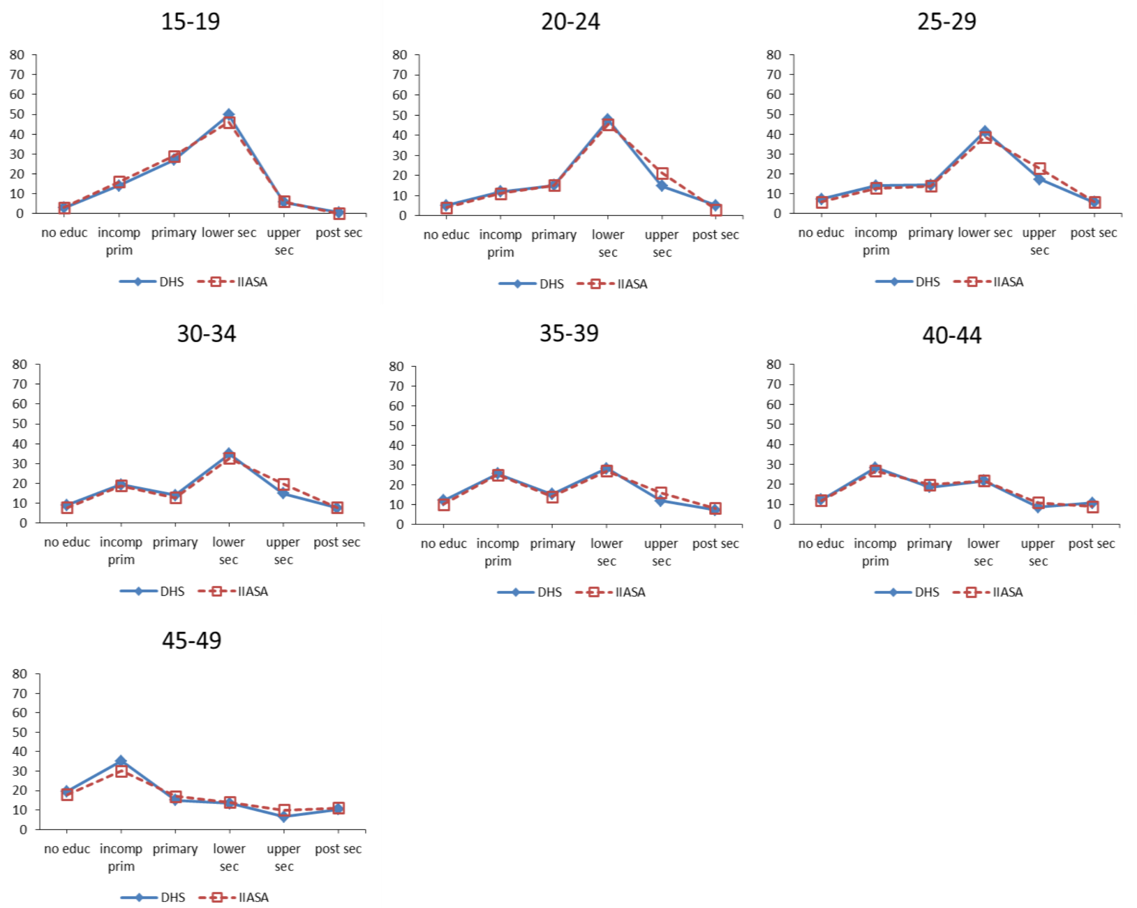


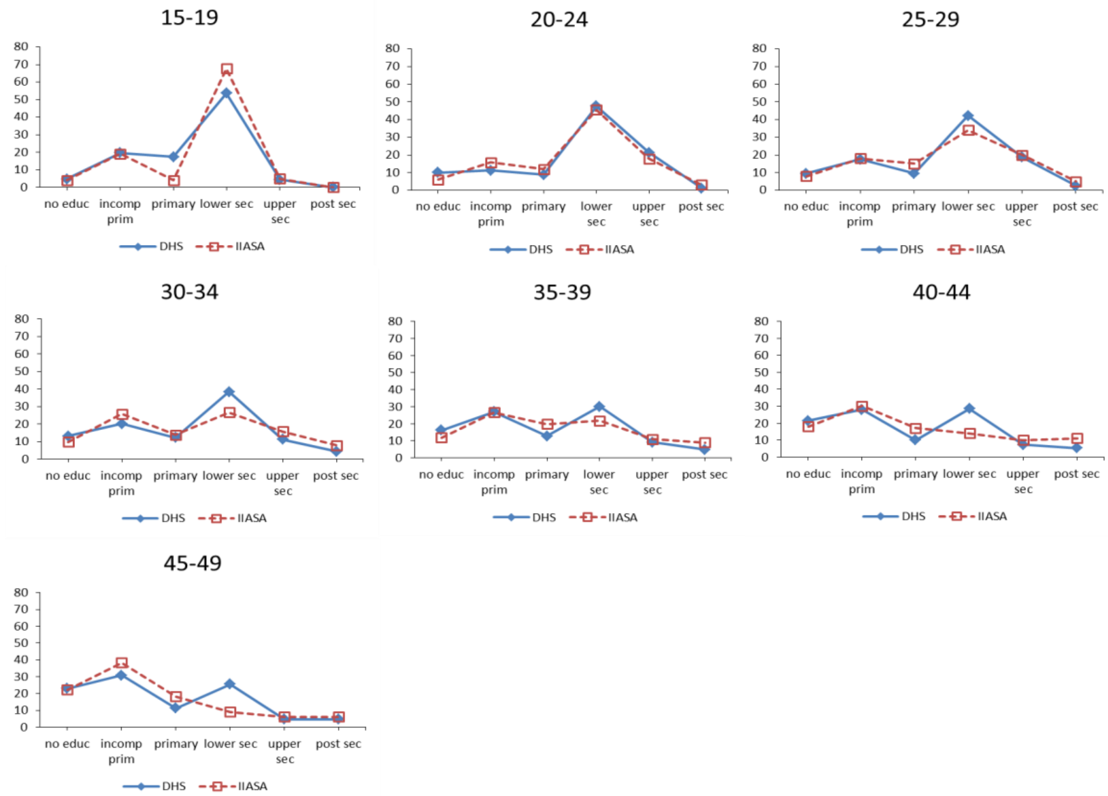

Figure 5-5  Percentage education distribution, Namibia 2013 DHS



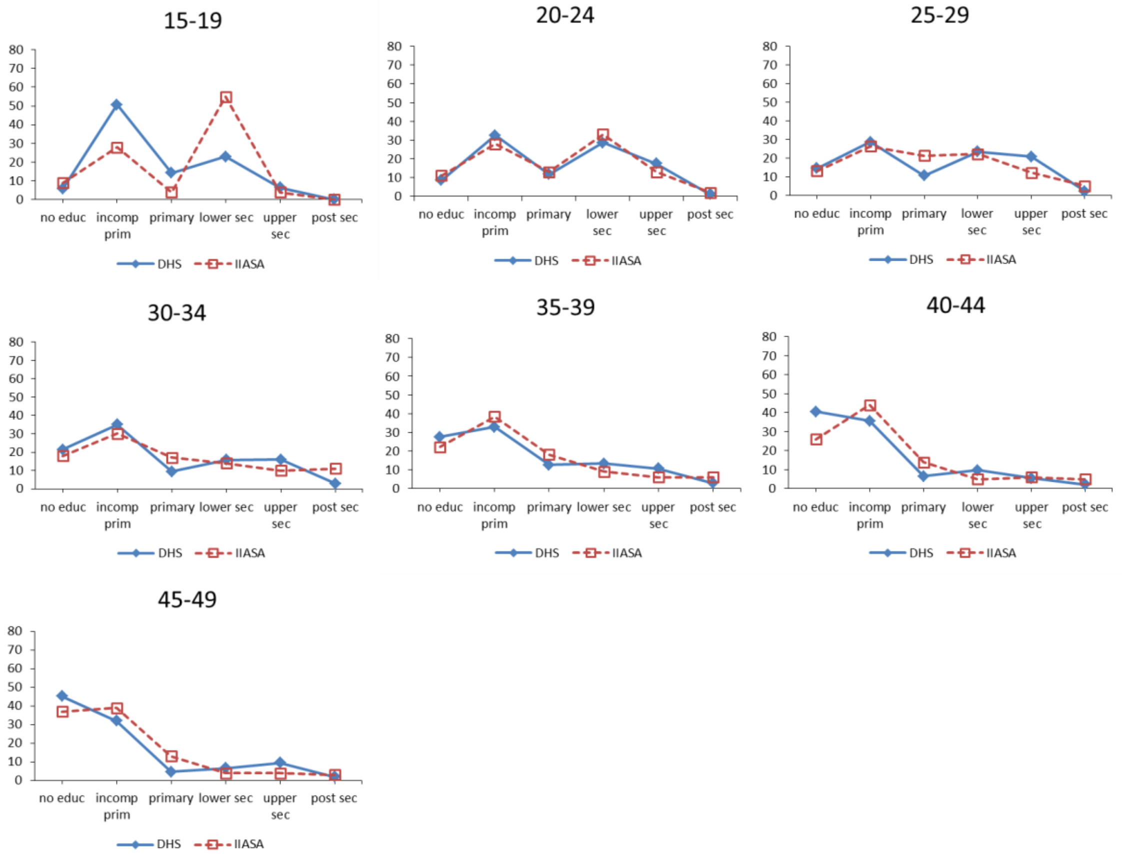
## Figure 5-6 Percentage education distribution, Namibia 2006-07 DHS




## Figure 5-7 Percentage education distribution, Namibia 2000 DHS



## Figure 5-8 Percentage education distribution, Namibia 1992 DHS




## Figure 5-9 Percentage education distribution, Lesotho 2014 DHS

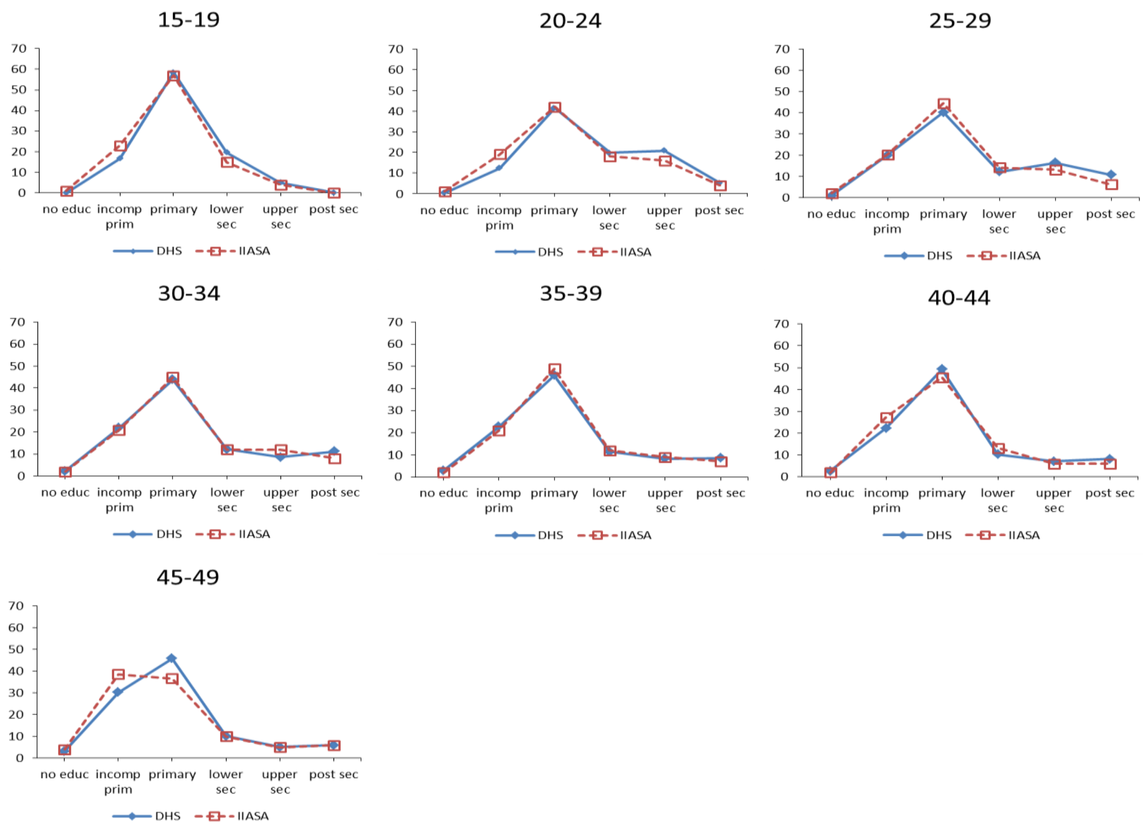


Figure 5-10  Percentage education distribution, Lesotho 2009 DHS

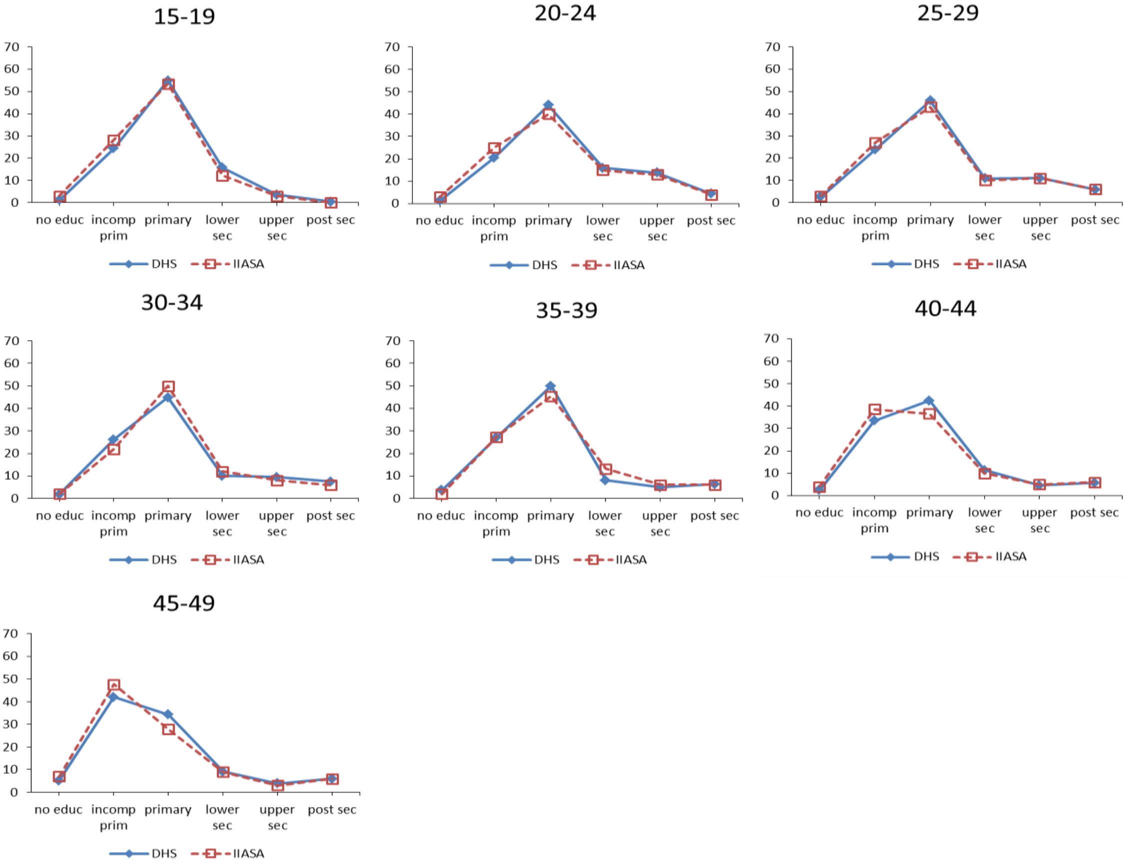
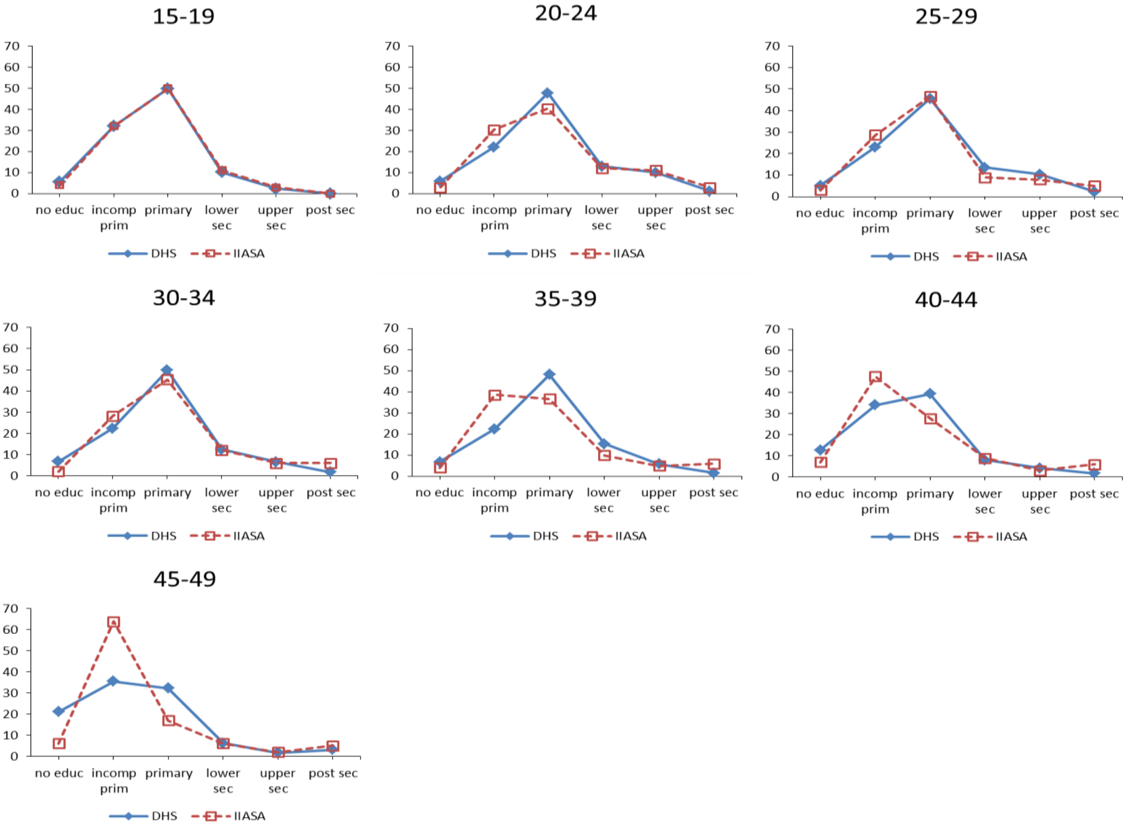
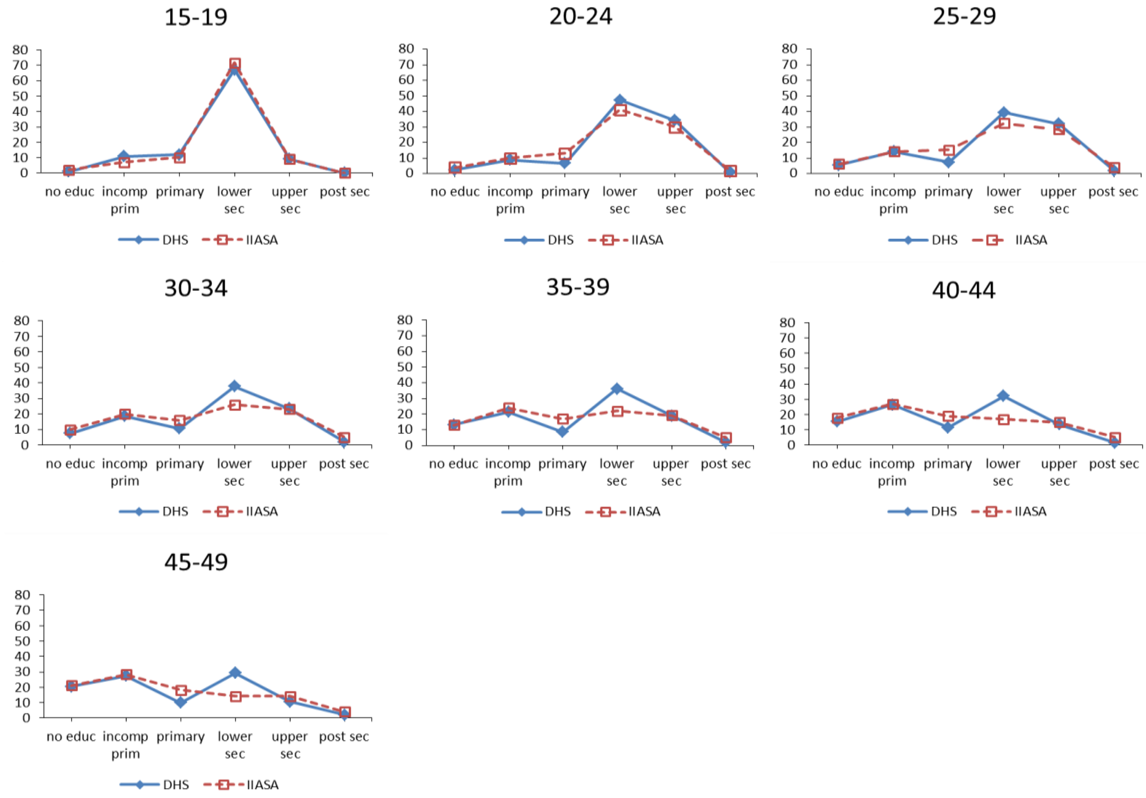


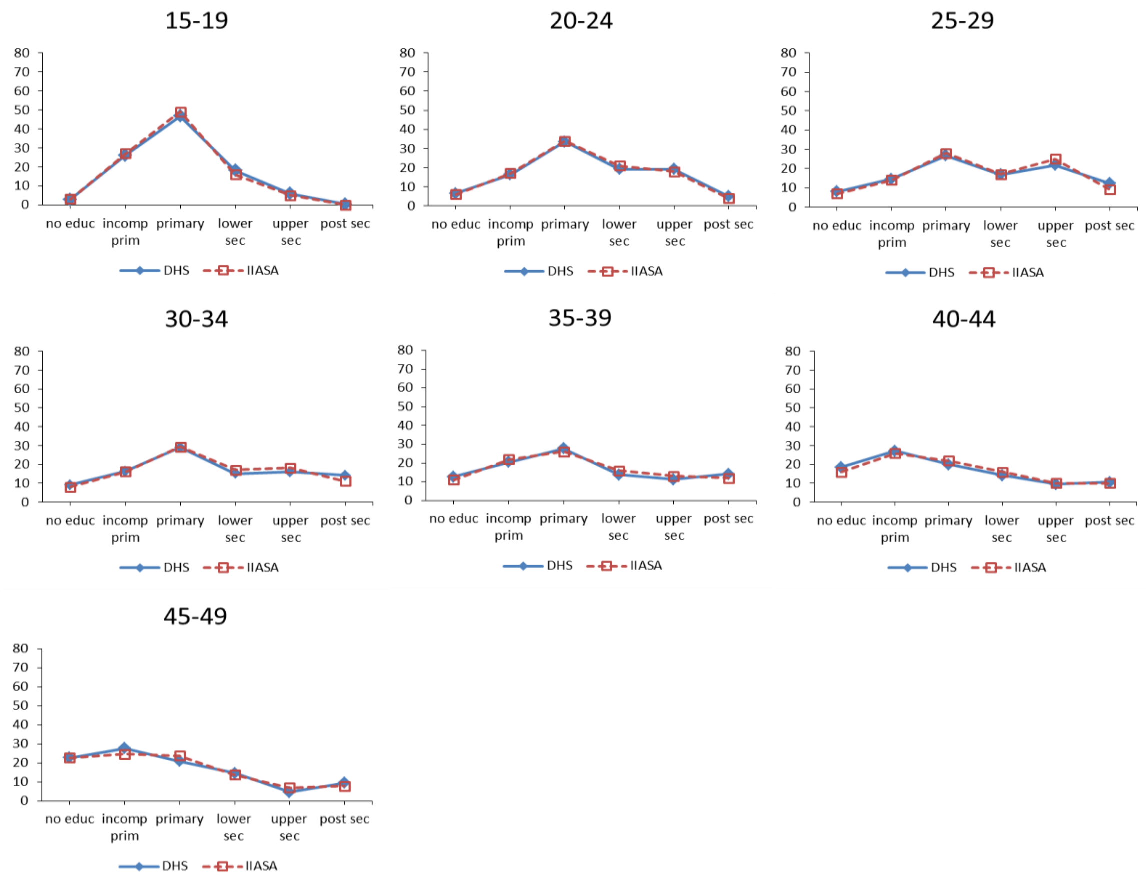

Figure 5-11  Percentage education distribution, Lesotho 2004 DHS



## Figure 5-12 Percentage education distribution, South Africa 1998 DHS

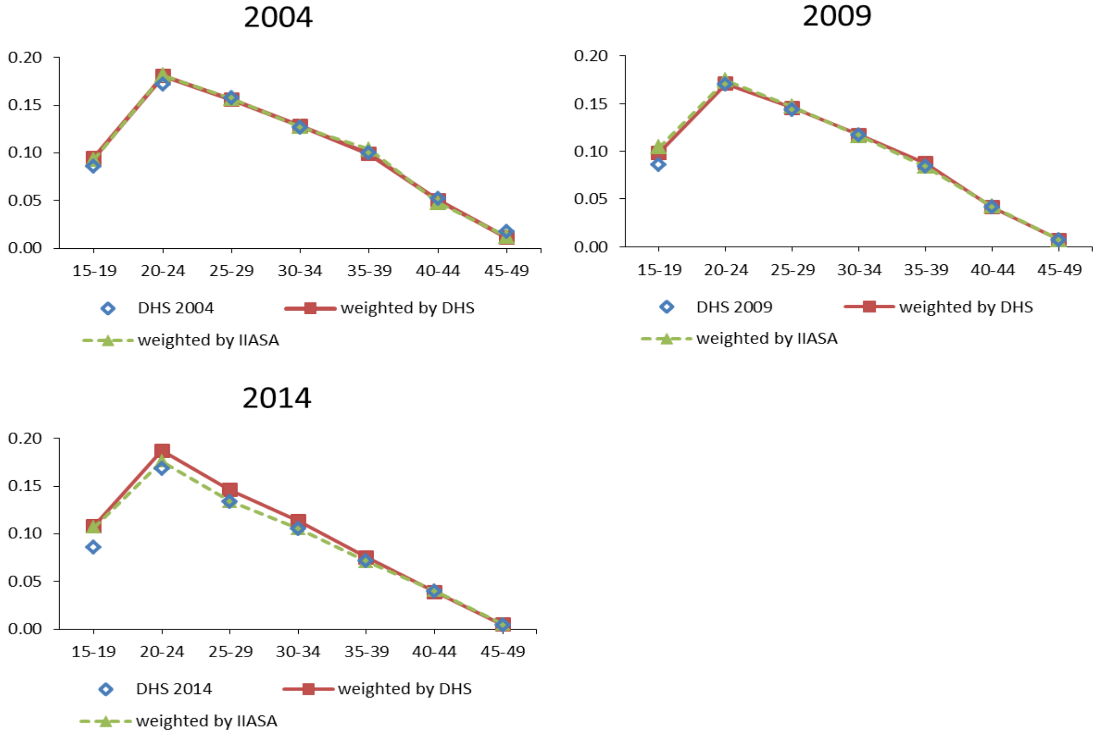



## Figure 5-13 Percentage education distribution, Swaziland 2006-07 DHS



# Appendix B National ASFRs for DHS and IIASA, Southern Africa

## Figure 5-14 National ASFRs in Lesotho




## Figure 5-15 National ASFRs in Namibia

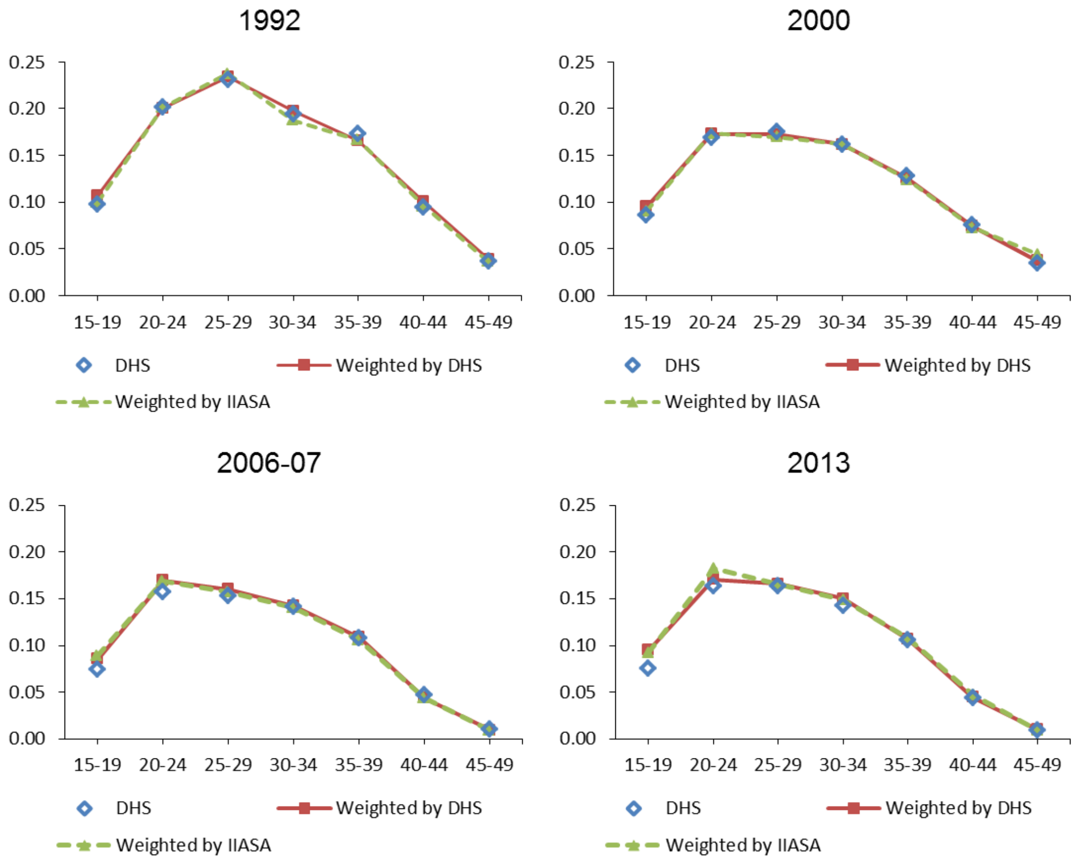


**Figure 5-16 National ASFRs in South Africa and Swaziland**

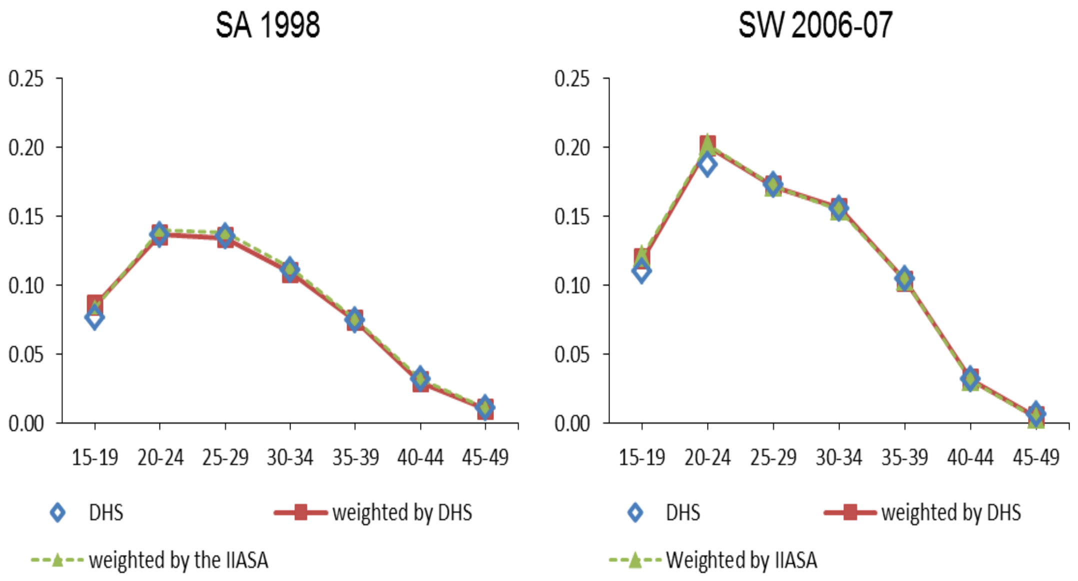



**Figure 5-17 National ASFRs in Zimbabwe**