Interrogation of the fertility differentials between the Malawi DHS and the Malawi Diffusion and Ideational Change Project survey data

Tapfuma Pashapa

Supervisor: A/Professor Tom Moultrie

University of Cape Town

Submitted in partial fulfilment of the requirements for a
Masters of Philosophy in Demography
at the Centre for Actuarial Research,
University of Cape Town

February 2011

PLAGIARISM DECLARATION

I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own. I have used the Harvard convention for citation and referencing. Each contribution to, and quotation in, this report from the work(s) of other people has been attributed, and has been cited and referenced. This research is my own work. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Signature: ______________________________

Date: __________________________________

ABSTRACT

Anglewicz, Adams, Obare *et al* (2009) show that the mean parities for the women who were interviewed in the Malawi Diffusion and Ideational Change Project (MDICP) surveys of 1998 and 2004 are generally higher than the mean parities for the women who were interviewed in the Malawi Demographic and Health Surveys (MDHSs) of 2000 and 2004 respectively. Our investigation is impelled by their concern that attrition and sample representativeness could contribute to bias when analysing the MDICP data. We interrogate the fertility differentials by four variables: age, marital status, contraceptive use, and education. The investigation shows that the comparisons by Anglewicz, Adams, Obare *et al* are confounded by age. We use the inter-survey parity increment method to estimate the fertility implied by the MDICP data. We find that the fertility rates of the women who were interviewed in the early MDICP surveys are inconsistent with the fertility of the women in the MDHSs. Thus, we conclude that the fertility rates which we observe for the early MDICP surveys are affected by errors in the data. Nevertheless, the results of our investigation suggest that the fertility of ever-married rural women declined marginally during the study period.

TABLE OF CONTENTS

PLAGIARISM DECLARATION 2

ABSTRACT 3

TABLE OF CONTENTS 4

LIST OF TABLES 6

LIST OF FIGURES 7

ACKNOWLEDGEMENTS 8

MAP OF MALAWI 9

1 INTRODUCTION 10

1.1 Background to study 10
1.2 Statement of research problem 11
1.3 Methodology 11
1.4 Thesis outline 12

2 LITERATURE REVIEW 13

2.1 Malawi Demographic and Health Surveys 13
2.2 Proximate determinants of fertility as they relate to Malawi's fertility 16
2.3 Malawi Diffusion and Ideational Change Project (MDICP) 21

3 DATA ANALYSIS 24

3.1 Investigation of MDICP data 24
3.2 Data from the MDHS, and comparison with data from MDICP 33
3.3 Mean parities of the MDICP and MDHS data 42

4 FERTILITY ESTIMATION 47

4.1 Fertility estimation method 47
4.2 MDICP fertility rates 50
4.3 Summary of results and the possible explanations 56

5 CONCLUSIONS 57

5.1 MDHS and MDICP sample differentials, and MDICP data quality ... 57

5.2 Findings 58
5.3 Limitations of our investigation 60
5.4 Conclusions and Recommendations 61

REFERENCES..63

APPENDIX...65

LIST OF TABLES

Table 2.1 Malawi fertility by background characteristics, 2000 and 2004 MDHS.........15
Table 2.2 Median age at first marriage, median age at birth and the percentage of premarital births for the women who were interviewed in the DHS's of Eastern and Southern African countries between 1990 and 2000...............18
Table 3.1 Mean parity and percentages for Malawi rural women, by family planning characteristics, 2000 MDHS, 2004 MDHS, 1998 MDICP, and 2004 MDICP24
Table 3.2 Distributions of women by age, marital status, contraceptive use, and education, 1998 MDICP, 2001 MDICP, 2004 MDICP, 2006 MDICP, and 2008 MDICP (Percentage distributions are given in brackets) 26
Table 3.3 Proportions of consistent ages before imputing missing ages, 1998 MDICP, 2001 MDICP, 2004 MDICP, 2006 MDICP, and 2008 MDICP......................28
Table 3.4 Distributions of differences between estimated and observed ages, by period (Percentage distributions are given in brackets) 28
Table 3.5 Proportions of ever-married rural women by age group and survey, 2004 MDICP, 2006 MDICP, and 2008 MDICP 30
Table 3.6 Proportions of ever-married women who ever used modern or traditional contraceptives, by age group and survey 31
Table 3.7 Proportions of ever-married women aged 15-49 who ever used modern or traditional contraceptives and were interviewed in all previous survey rounds, by age group and survey 31
Table 3.8 Proportions of ever-married women with some education, by age group and survey 32
Table 3.9 Proportions of ever-married women with some education, who were interviewed in all previous survey rounds, by age group and survey 32
Table 3.10 Proportions of women living in urban areas by age group and survey, 2000 MDHS and 2004 MDHS...33
Table 3.11 Proportions of ever-married rural women by age group and survey 38
Table 3.12 Proportions of ever-married rural women who have ever used contraceptives, by age group and survey 39
Table 3.13 Proportions of women aged 15-49 who have some education, by age group and survey...40
Table 3.14 Proportions of ever-married rural women aged 15-49 who have some education, by age group and survey 40
Table 3.15 Distributions of parity increment by MDICP inter-survey period (Percentage distributions are given in brackets)..43
Table 3.16 Mean parities of ever-married rural women aged 15-49, by age group and survey 43
Table 4.1 Observed MDICP marital fertility rates by five year age group and inter-survey period 50
Table 4.2 Comparison of the Observed and Brass fitted marital fertility rates for the 2004-2008 MDICP inter-survey period to rates for the women who were interviewed in the 2000 MDHS..55

TABLE IN APPENDIX

Table A.1 Observed and Brass fitted age specific fertility rates, 2004-2008 MDICP inter-survey period...65

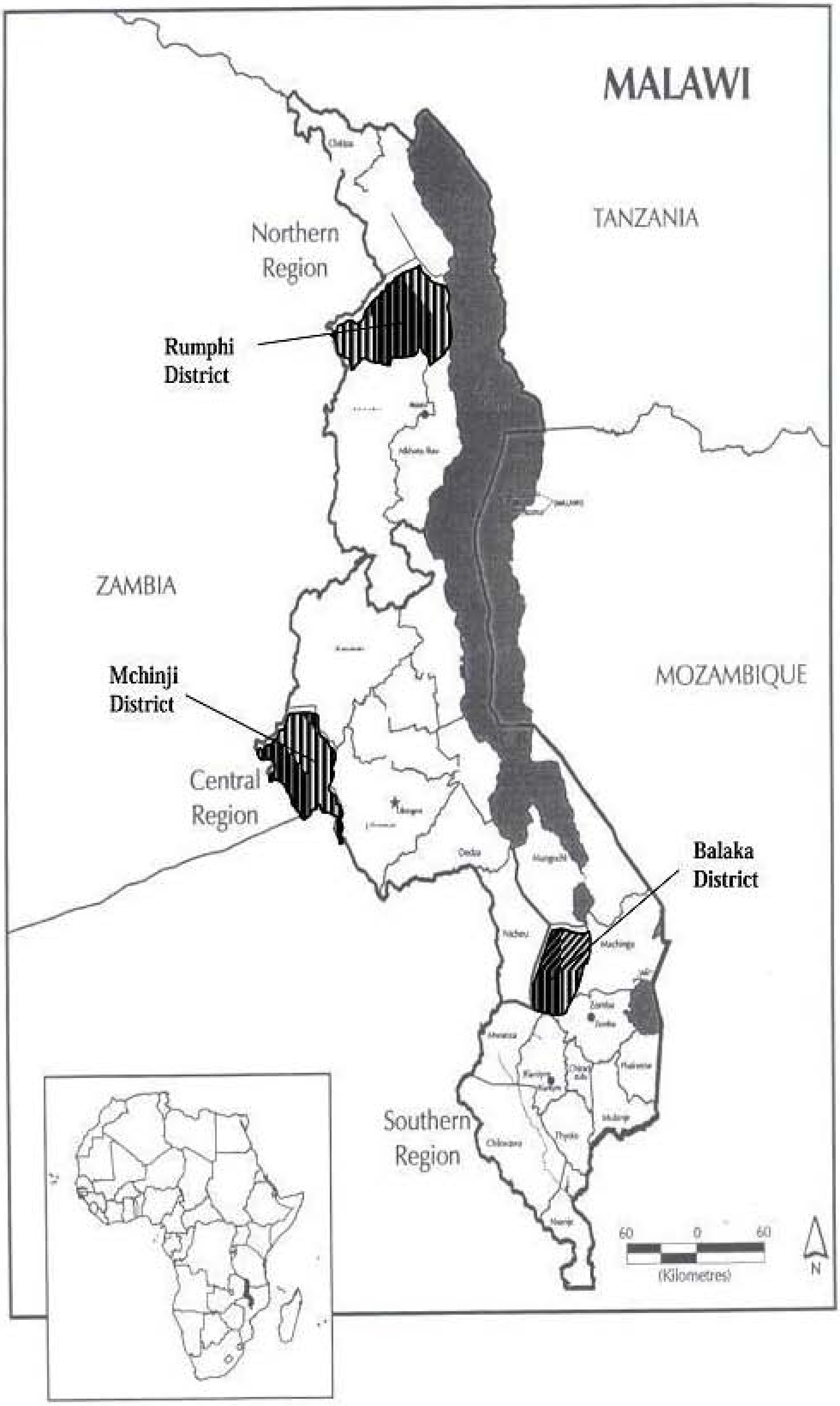
LIST OF FIGURES

Figure 2.1 Malawi total fertility rates from 1984 to 2004 14
Figure 3.1 Age distributions of women aged 15-49, who were interviewed in the MDICP surveys, 1998 MDICP, 2001 MDICP, 2004 MDICP, 2006 MDICP, and 2008 MDICP 29
Figure 3.2 Age distributions of ever-married rural women aged 15-49, 2000 MDHS and 1998 MDICP..34
Figure 3.3 Age distributions of ever-married rural women aged 15-49, 2000 MDHS and 2001 MDICP..35
Figure 3.4 Age distributions of rural women aged 15-49, 2004 MDHS and 2004 MDICP 36
Figure 3.5 Age distributions of all the rural women aged 15-49, 2004 MDHS and 2006 MDICP 37
Figure 3.6 Age distributions of all the rural women aged 15-49, 2004 MDHS and 2008 MDICP 37
Figure 4.1 Observed MDICP inter-survey marital fertility rates, 1998 to 2001, 2001 to 2004, 2004 to 2006, and 2006 to 2008..51
Figure 4.2 Observed marital fertility rates of the MDICP inter-survey periods, 2004-2006 and 2006-2008 54
Figure 4.3 Comparison of the Observed and Brass fitted marital fertility rates for the 2004-2008 MDICP inter-survey period to the rates for women who were interviewed in the 2000 MDHS..55

ACKNOWLEDGEMENTS

I would like to thank Professor Tom Moultrie who supervised this research. I also want to thank Andrew W Mellow Foundation for financing my studies at University of Cape Town. I am grateful to my family and friends for their encouragement and support throughout the years.

MAP OF MALAWI



Source: Watkins, Zulu, Kohler *et al* (2003)

1 INTRODUCTION

Previous research on fertility using the Malawi Demographic and Health Survey (MDHS) data has showed, as in other sub-Saharan African countries, that rural women had higher fertility than urban women, and they also lagged behind in terms of HIV/AIDS awareness. The gap in HIV/AIDS awareness provided the essential motivation for the establishment of the Malawi Diffusion and Ideational Change Project (MDICP). The MDICP is a "longitudinal research project with the overall goals of investigating the multiple processes and influences that contribute to variation in HIV risks in a sub-Saharan Africa context, identifying risks and assessing the potential effect of HIV risk reduction programs on infection risks and disease dynamics" Anglewicz, Adams, Obare *et al* (2009: 504). In this context, the research which has been done on MDICP data is mainly about investigating the effects of social networks on demographic behaviour and perceptions regarding HIV/AIDS. The need to measure the fertility implied by the MDICP data arises when studying the relationship between HIV/AIDS awareness and the change in individual fertility preferences. However, such a study may produce misleading results because of the errors which are generally found in the data of longitudinal surveys. Therefore, in order to determine the quality of the MDICP fertility data as well as the extent to which the data may offer insight into national patterns, we are going to investigate the MDHS and MDICP data sets by the determinants of fertility and then compare the fertility of the women who were interviewed in the two surveys.

1.1 Background to study

Our investigation is an extension of the analysis of the 1998 and 2004 MDICP data sets by Anglewicz, Adams, Obare *et al* (2009). They investigated four potential sources of bias: sample representativeness, interviewer effects, response unreliability, and sample attrition. These aspects became important because the MDICP was in its fifth round. Their main concern was the attrition of the initial cohorts of women who were interviewed in 1998, which had accumulated over time.

Anglewicz, Adams, Obare *et al* say that the characteristics (age, socio-demographic and HIV/AIDS-related, and fertility and family planning characteristics) of the MDICP respondents differ significantly from those of the MDHS respondents. They suspect that the differentials are caused by the fact that the respondents who were interviewed in the two surveys were sampled from two different subsets of the rural population or by the addition of a new sample of respondents in 2004. Moreover, Reniers (2003) had found that the women who were lost to attrition between the MDICP surveys of 1998 and 2001 had fewer children, and were less likely to be members of indigenous churches compared to those who were interviewed in both surveys.

Anglewicz, Adams, Obare *et al* calculated the mean parities for the women who were interviewed in the MDICP surveys of 1998 and 2004 and then compared them with the parities for the women who were interviewed in the 2000 and 2004 MDHS. Their results show that the mean parities for the women who were interviewed in the MDICP surveys are generally higher than the parities for the women who were interviewed in the MDHS. We suspect that the differentials are caused by age confounding.

Thus, we are going to test the hypotheses that the mean parities which were compared by Anglewicz, Adams, Obare *et al* are affected by age confounding as well as other factors governing fertility.

1.2 Statement of research problem

Our investigation is an extension of the analysis by Anglewicz, Adams, Obare *et al* in which they investigate the sources of bias in the MDICP data sets. However, our investigation is focused on fertility data. Our concern is that the mean parities which are presented by Anglewicz, Adams, Obare *et al* for the MDICP data sets of 1998 and 2004 are generally higher than the mean parities for the 2000 and 2004 MDHS data sets. Their comparison does not provide any supporting evidence for the observed differentials. Thus, we want to determine the reasons for the fertility differentials between the MDHS and MDICP data sets. We shall investigate the differentials with respect to the determinants of fertility.

1.3 Methodology

First, we will investigate the age and parity data for consistency, and then we impute some of the missing data. We shall investigate the distribution of the women by age and the other determinants of fertility. Our aim is to determine the effects of the determinants on the fertility of the women who were interviewed in the MDICP surveys. In order to estimate fertility from the MDICP data, we apply a variation of the inter-survey parity increment method as set out in United Nations (1983). We shall use the parity data in each data set to determine the number of births by each woman during the period between consecutive surveys. We shall smooth the age distributions of our fertility estimates by using the Brass polynomial.

1.4 Thesis outline

Our investigation will begin by reviewing the literature on fertility in Malawi and sub-Saharan Africa. In chapter three, we analyse the MDICP data sets in order to establish how they differ from the MDHS data sets. In chapter four, we estimate the fertility rates of the women who were interviewed in the MDICP surveys and compare them with the fertility of the women who were interviewed in the MDHSs. At the end of our investigation we will have gained an understanding of the MDICP fertility data. Hence, we present our findings and recommendations in chapter five.

2 LITERATURE REVIEW

In this chapter, we assess the literature about fertility in Malawi. First, we present the fertility trend as it is observed in the Malawi Demographic and Health Survey (MDHS) reports. Then, we review the findings of earlier studies on fertility in Malawi. We conclude the chapter by presenting the Malawi Diffusion and Ideational Change Project (MDICP).

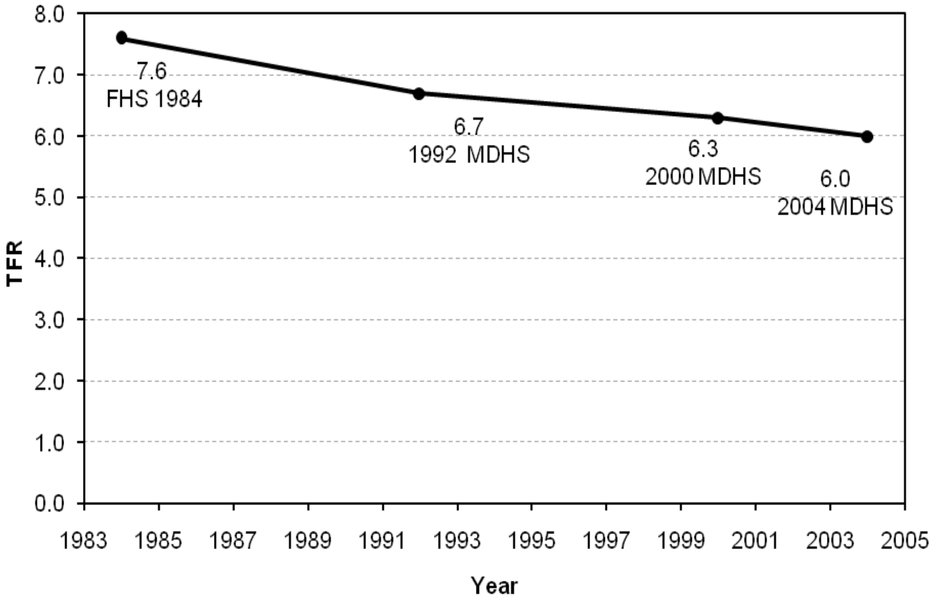
2.1 Malawi Demographic and Health Surveys

The main sources of fertility data for Malawi are the MDHSs. The surveys are designed to provide estimates of health and demographic indicators at national and regional levels for rural and urban areas. MDHSs are held regularly and they follow the same standard procedures of Demographic and Health Surveys that are conducted in other African countries. The MDHS survey design is a clustered multistage sample which is drawn from the Malawi census frame. Before each survey, cluster samples are drawn from the census sample frame and then an exhaustive listing of the households in each cluster is done. Households are then randomly sampled from each cluster and in the sampled households; all the women who are aged 15-49 years are eligible for individual survey interviews. According to the 2004 MDHS report by National Statistical Office [Malawi] and ORC Macro International (2005), the National Statistical Office of Malawi has the responsibility of conducting the surveys, while technical support is provided by ORC Macro International. The Ministry of Health and Population, the National AIDS Commission (NAC), the National Economic Council, and the Ministry of Gender also contribute to the development of questionnaires.

2.1.1 Fertility trend according to MDHS reports

Beginning 1984 to 2004, the total fertility rate (TFR) of Malawi declined by 21 per cent. The decline is more clearly explained by Mijoni (2005). He used cohort-period fertility rates and parity progression ratios on the 2000 MDHS data to show that lower proportions of women by age groups proceeded to a higher order birth within a five year period than observed in the 1992 MDHS. In addition, the projected median birth intervals for the 2000 MDHS were generally longer than those for the 1992 MDHS. Therefore, this showed that there was a declining trend of a preference for higher order parities. Figure 2.1 shows the fertility trend of Malawi from 1984 to 2004. The fertility rates were obtained from the 2004 MDHS report by the National Statistical Office [Malawi] and ORC Macro International (2005).

Figure 2.1 Malawi total fertility rates from 1984 to 2004



Source: National Statistical Office [Malawi] and ORC Macro International (2005)
Note: FHS: Family Formation Survey, MDHS: Malawi Demographic and Health Survey

The fertility among rural women who were interviewed in the 2004 MDHS was 6.4 children per woman, while the fertility among urban women was 4.2 children per woman. Both rates had declined marginally from the rates for the women who were interviewed in the 2000 MDHS. The urban-rural comparison of the TFR shows a difference of more than one child per woman in both surveys. The urban-rural differentials for the women who were interviewed in the 2000 MDHS were also explored by Mijoni. He explains that the proportions of urban women who proceeded to a next birth were lower than those of their counterparts in rural areas. Mijoni used parity progression ratios to show that, although the parity progression ratios for the 2000 MDHS data indicated that fertility was declining, the ratios for women in rural areas were constant. This showed that there was a slower pace of fertility decline for women who lived in rural areas. Table 2.1 summarises the characteristics of the women who were interviewed in the 2000 MDHS and the 2004 MDHS.

In 2004, the TFR for the Central region was 6.4 children per woman, while in the Southern and Northern regions it was 5.8 and 5.6 children per woman respectively. From 2000 to 2004, the provincial TFR decline was smaller for the Central region compared to the other regions.

The median age at first marriage for women in the 20-49 age range in 2000 and 2004 was approximately 18 years. In 2004, 51 per cent of the women got married before 18 years. The median age at first intercourse for women increased from 16.9 years in 2000 to 17.3 years in 2004.

The education variable is the most important variable when analysing fertility by the women's background characteristics. We also note that the TFR for women in the poorest wealth quartile was 7.1 children per woman compared to a rate of 4.1 children for women in the richest quartile of all the women who were interviewed in the 2004 MDHS.

Table 2.1 Malawi fertility by background characteristics, 2000 and 2004 MDHS

Characteristic	*2000 MDHS*	*2004 MDHS*
Age in years		
*Age at first intercourse	16.9	17.3
*Age at first marriage	17.9	18.0
*Age at first birth	19.1	19.0
Median months since preceding birth	33.8	35.9
TFR by education		
No education	7.3	6.9
Primary (1 to 4 years)	6.7	6.6
Primary (5 to 8 years)	6.0	5.8
Secondary and higher education	3.0	3.8
TFR by wealth quartile		
Lowest	-	7.1
Second	-	7.0
Middle	-	6.5
Fourth	-	5.8
Highest	-	4.1
TFR by region		
Northern	6.2	5.6
Central	6.8	6.4
Southern	6.0	5.8
TFR by residence		
Urban	4.5	4.2
Rural	6.7	6.4
Age Specific Fertility Rates by age group		
15-19	0.172	0.160
20-24	0.305	0.291
25-29	0.272	0.252
30-34	0.219	0.222
35-39	0.167	0.162
40-44	0.094	0.088
45-49	0.041	0.038
Total Fertility Rate	6.3	6.0

Source: National Statistical Office [Malawi] and ORC Macro international (2001) and (2005)
Note: *For women aged from 20 to 49 years.

Fertility decline such as it has been in Malawi can be attributed to the introduction of favourable family planning policies as explained by Chimbwete, Watkins and Zulu

(2005). We also discuss the other determinants of fertility as they relate to fertility in Malawi.

2.2 Proximate determinants of fertility as they relate to Malawi's fertility

Bongaarts (1984) classifies the proximate determinants of fertility into two categories: proximate variables, and socioeconomic and environmental variables. He says that the latter variables include the social, cultural, economic, institutional, psychological, health, and environmental variables, while the proximate variables consist of all the biological and behavioural factors through which the socioeconomic and environmental variables must operate to affect fertility as stated in an earlier publication by Bongaarts and Potter (1983). Some of the proximate variables which he lists are; the proportion of married women or in sexual union, frequency of sexual intercourse, postpartum abstinence, induced abortion, and contraception.

2.2.1 Family planning and contraception

Contraceptive use is important because it limits the number of unwanted pregnancies and it controls the timing of births. Garenne (2008) regards the increased use of modern contraceptives as the leading driver of fertility decline in sub-Saharan Africa because of the success of different reproductive health programs and policies.

Khan, Shane, Mishra *et al* (2007) conducted a research on the trend of contraceptive use in developing countries. In these countries, contraceptive knowledge was nearly universal, irrespective of place of residence and educational level. Although most women were aware of multiple methods of contraception in the countries studied, this awareness was lower in sub-Saharan Africa than in all the other regions of the world. In addition, the gap between the knowledge of contraception and ever use of contraception was larger in sub-Saharan Africa than anywhere else. Khan, Shane, Mishra *et al* state that spousal discussion of family planning and the knowledge of multiple methods of contraception were also low in the region. Despite the increase of contraceptive use in most African countries, they assert that special effort was needed to reach rural areas and less educated women because they continued to have low levels of contraceptive use.

In Malawi's case, Chimbwete, Watkins and Zulu (2005) explain that, under Kamuzu Banda's governance from 1966 to 1994, modern family planning methods were not allowed because they were considered to be un-African. The Banda administration also disputed the idea that population growth was a hindrance to development. In addition, Malawi's development policy under Banda's rule was based on subsistence agriculture as a means of livelihood. In Africa, subsistence agriculture is associated with the idea of having a large family in order to provide a large labour force to work in the family's fields. Therefore, Banda's policies contributed to Malawi's high fertility.

According to Chimbwete, Watkins and Zulu, in 1980, the Malawi government only allowed the use of traditional methods for child spacing after being pressured by non-governmental organisations to change the family planning policy. The policy was then named; Malawi's Child Spacing Policy. Family planning was fully introduced in 1994 after a change in government, when the policy was renamed the Malawi Family Planning Program (MFPP). During the same year, at the Cairo Population Conference of 1994, a resolution on reproductive rights was passed. This meant that the MFPP had to be revised by incorporating reproductive rights so that it could receive funding from international donors.

Chimbwete, Watkins and Zulu explain that, in 2001, United Nations Population Fund (UNFPA) and the Department of Population Services of Malawi formed a multisectoral committee to revise Malawi's population policy in order for it to adhere to the Cairo Programme of Action. According to Solo, Jacobstein and Malema (2005), one of the notable changes brought by the new policy occurred in 2002 when the United Kingdom Department of International Development provided funds for a non-governmental organisation (NGO) called Banja La Mtsogolo (BLM). BLM was formed in 1987 and it aims to achieve better sexual and reproductive health and to address HIV/AIDS—especially among people who are poor and hard to reach. According to a report by Banja La Mtsogolo (2008), the organisation receives management support from Marie Stopes International and it has 30 centres across the country. BLM chairs the Sexual and Reproductive Health NGOs in Malawi. The organisation also collaborates with the Family Planning Association of Malawi (FPAM), the Girl Guides Association (MAGGA), the National Youth Council in Malawi, and the Council for Non-Governmental Organisations in Malawi (CONGOMA). According to the report, BLM is apolitical and it works with traditional and political leaders at all levels. Solo, Jacobstein and Malema explain that, between 1992 and 2000, the progress in the use of modern contraceptives caused the contraceptive prevalence rate (CPR) to increase from 7.4 to 26.1 per cent among married women aged 15-49 years.

2.2.2 Marital status

Harwood-Lejeune (2000) examined the trends of fertility and marriage in nine countries in Southern and Eastern Africa by using the Demographic and Health Survey (DHS) data of the surveys which were conducted between 1990 and 2000. She discovered that the countries with well-established fertility declines like Kenya and Zimbabwe had a higher age at marriage, a higher age at first birth, and a higher percentage of premarital births than the countries which had no evidence of fertility decline like Uganda. In most of these countries, between 20 and 30 per cent of the first births were premarital. Table 2.2 shows the median age at first marriage, median age at birth, and the percentage of premarital births for the women who were interviewed in the DHS's.

Harwood-Lejeune concludes that rising age at first marriage is one of the factors driving the fertility declines of seven out of the nine countries which she studied, and that approximately one sixth to one third of the fertility declines among women aged 15-39 years is explained by the rising age at first marriage. These findings show that marriage is an important factor in determining the levels of fertility because they suggest that, in the countries studied, most women give birth after their first marriage. In Malawi's case, Garenne (2008) says that the proportion of premarital births averaged around three per cent from 1960 to 2004, as indicated by the data of early censuses and recent MDHSs. Despite this, Malawi has one of the highest rates of divorce in the world.

Table 2.2 Median age at first marriage, median age at birth and the percentage of premarital births for the women who were interviewed in the DHS's of Eastern and Southern African countries between 1990 and 2000

Country (Year of DHS)	*Median age at first marriage*	*Median age at first birth*	*Interval between median ages at marriage and first birth*	*Percentage of first births which are premarital*
Uganda (1995)	17.75	18.83	1.08	20
Mozambique (1997)	17.75	19.17	1.42	19
Malawi (1992)	17.92	19.00	1.08	13
Zambia (1996)	18.08	18.75	0.67	22
Tanzania (1991-1992)	18.08	19.00	0.92	21
Tanzania (1996)	18.58	19.33	0.75	21
Madagascar (1997)	18.75	19.50	0.75	23
Zimbabwe (1994)	19.33	19.75	0.42	24
Kenya (1998)	19.67	19.75	0.08	32
Namibia (1992)	24.42	20.75	-3.67	56

Source: Harwood-Lejeune (2000)

Reniers (2003) asserts that, in Malawi, marriage is a fragile institution and in Balaka district of southern Malawi, one third of the marriages do not last beyond the fifth anniversary. According to Reniers, marital instability is mainly caused by labour migration as males separate from their spouses. In addition, divorces are high in marriages in which women are employed. He argues that the high rates of divorce stimulate high rates of remarriage. Therefore, remarriage minimises the effect of divorce on fertility. Reniers also notes that polygamy is higher in the districts in which HIV prevalence is low. He suggests that this is caused by the fact that polygamy reduces the number of extramarital affairs and hence the spread of HIV/AIDS.

In another study of the fertility of the Agincourt district in South Africa, Garenne, Tollman and Kahn (2000) say that marriage was increasingly being delayed by women, and the proportions of never-married women were increasing at all the ages because of the erosion of traditional ways. Thus, they note that the age specific fertility distribution is bimodal. The first mode is for premarital fertility (among women aged 12-26 years), and the second mode is for marital fertility (among women aged 28-35 years).Their concern is that the first mode is caused by a low incidence of contraceptive use before the first birth (especially among adolescents), a low prevalence of abortion, and a high contraceptive prevalence thereafter. Garenne, Tollman and Kahn suspect that this age specific fertility pattern describes the age distribution of the fertility of most rural societies in other African countries.

2.2.3 Postpartum abstinence

Zulu (2001) conducted a study on postpartum abstinence by using the data from the 1988 Traditional Methods of Child Spacing in Malawi (TMCSM) survey and the 1998 MDICP survey. In both surveys, women were asked to state how many months had passed before they terminated breast feeding, began menstruation, and resumed sexual intercourse after their last birth in the preceding five years.

The studies showed that the duration of postpartum abstinence did not change substantially between 1988 and 1998. Further, six months after giving birth, 88 per cent of the women in the Southern region, 94 per cent of the women in the North, and 34 per cent of women in the Central region still abstained from sex. After 24 months, 30 per cent abstained in the North as compared to proportions of less than 10 per cent in the South and Central Regions.

Zulu says that in Malawi, sexual abstinence after child birth was viewed as a way of protecting the mother's milk from contamination by semen, and it was also meant to protect the health of the man. He explains that throughout the country, the end of postpartum bleeding was taken to mean that a woman had "cooled down" by eliminating all of the "dangerous blood" in her body after child birth. In all the three regions, the exact duration of abstinence was not specified by the societies. However, Zulu's study showed that a ritual to mark the end of postpartum abstinence was held around the time when a woman's child started sitting in the Central, crawling in the Southern, and just after the commencement of menstruation in the Northern region. These studies showed that most women who abstained from sex after a birth did not know about the contraceptive effect of this period of abstinence.

By constructing life tables on the probabilities of abstinence and on the resumption of menstruation, Zulu noted that the survival probabilities for abstinence were higher than those for the resumption of menstruation in the Northern region. This showed that women in the Northern region were more likely to abstain from sex after birth, and that they abstained until after the resumption of menstruation. A converse pattern was found for women who lived in the other regions. In 1988, the mean abstinence times for women in the Central, Southern and Northern regions were 7, 10 and 16 months and in 1998, they were 7, 10 and 17 months respectively.

Zulu concludes that postpartum abstinences is not a means of contraception but rather that traditional methods of contraception such as medicines that women drink, charms that they wear around the waist, reduced frequency of sex, coitus interruptus and periodic abstinences are. In sum, since the majority of the women in the country abstained from sex for a period which ended before the resumption of menstruation, the practice of postpartum abstinence was not used to control fertility.

2.2.4 Education

Education is a socioeconomic fertility determinant because of the tendency by educated women to use contraceptives and to marry at an older age. Kravdal (2001) says that fertility is influenced by individual women's education because of the high opportunity costs of childbearing involved in some types of work that may be offered to the better educated women. Therefore, educated women will tend to delay or limit the number of children that they have during their life time.

Doctor (2005) examined the association between religious affiliation and women's schooling by using data from the Malawi DHS of 2000. After controlling for childhood residence, age, parity, and age at first marriage, his findings suggest that a women's schooling is strongly influenced by her urban childhood residence and an increase in age at first marriage. This notion is supported by Harwood-Lejeune (2000) in her study of fertility in Southern and Eastern African countries. She asserts that educated and urban women have an older age at first marriage and birth. In addition, educated women tend to have more knowledge about contraceptives than those who are uneducated. Doctor also finds that, there are differentials in the acquisition of schooling by religious affiliation. He shows that for all age cohorts, women from the Church of Central Africa Presbyterian (CCAP) and other Christian denominations are more likely to have some schooling compared to women who belonged to other religions.

In another study, Garenne (2008) also found that, in sub-Saharan Africa, fertility decline occurred in areas with low levels of education, and that minimum education levels were necessary for this decline.

2.2.5 Age

An investigation of the age distribution is important because women of the same age group generally have the same fertility characteristics. Therefore, if women of the same age constitute a large proportion of the sample, then the fertility of the group will have a large weight on the overall fertility of the sample. This is of great concern when fertility is estimated by using the measures of central tendency because they are affected by confounding. Hence, age is an important fertility determinant because of the ages which are associated with the period of each woman's life when certain biological fertility determinants are effective.

According to Bongaarts (1984), a small proportion of women are sterile at the beginning of the reproductive years and this proportion increases after 50 years. Therefore, the measures of central tendency can be affected by the proportion of women in these sterile age groups. Further, marriage is also associated with age because women in rural African societies generally marry just before the age of 20 years.

Another concern is that if there are very few women in an age group then the estimated age specific mean parity may be inflated by small data changes. This is a common source of error when estimating the fertility of women in the oldest and youngest age groups because their fertility is usually very low. Another common error is caused by the preference for even numbered ages by women who do not know their ages. This causes age heaping and results in biased fertility estimates.

2.3 Malawi Diffusion and Ideational Change Project (MDICP)

The MDICP is an ongoing longitudinal study which was first conducted in 1998. Subsequent survey rounds were held in 2001, 2004, 2006, and 2008. According to Anglewicz, Adams, Obare *et al* (2009), one of the objectives of the MDICP is to investigate the role of social interactions in changing demographic attitudes and behaviour in rural Malawi.

The MDICP surveys cover villages in Balaka in the Southern, Mchinji in the Central, and Rumphi in the Northern regions of the country. Rumphi is inhibited by the

Tumbuka who are predominately protestant. The Tumbuka people follow the patrilineal system of kinship and lineage where residence is ideally patrilocal. Balaka district follows a matrilineal system of kinship and lineage system where residence is ideally matrilocal. Balaka district is inhibited by the Yao who are predominantly Muslim. Mchinji distict is inhibited by the Chewa who are predominantly Catholics and Protestants.

According to Watkins, Zulu, Kohler, and Behrman (2003), the three sites were chosen in order to carry out comparisons with the earlier surveys which were conducted in each of the respective districts. They explain that Mchinji (central region) and Rumphi (northern region) districts were designed to cover census enumeration areas covered in the TMCSM survey of 1988. However, the two MDICP samples differed from the TMCSM samples because the TMCSM samples consisted of both ever-married and never-married women. Hence, the Census Enumeration Areas which were sampled during the 1988 TMCSM included a smaller proportion of ever-married women than the MDICP samples. According to Zulu (2001), the 1998 TMCSM sampled respondents from Chiradzulu district in the southern region but Balaka district was instead selected for the MDICP study in order to take advantage of a family planning intervention program in the district. Miller, Zulu and Watkins (2001) explain that the Balaka sample was also designed in order to include the respondents who were covered in a survey that was conducted by a Germany organisation GTZ in 1993. Hence, the MDICP sample in the district was meant to measure the impact of a community based contraceptive distribution (CBD) program and so half of the respondents in the district were sampled from the intervention area and the other half from the non-intervention. Balaka was also chosen because it had a high proportion of Yao speaking Muslims compared to Chiradzulu. Watkins, Zulu, Kohler, and Behrman explain that the MDICP sampling strategy was not designed to be representative of the national population of Malawi, even though the sample characteristics closely match the characteristics of the rural population of the Malawi 1996 DHS.

The 1998 and 2001 surveys interviewed ever-married women only. The 2004 survey included some never-married adolescents aged 14-24 years in order to adjust for the gradual aging of the women and to include never-married women in the longitudinal study. According to Weinreb (2007), the 2004 survey only added young never-married women to allow for a prospective study of marriage among a cohort of young Malawi adults in the subsequent surveys. No changes were made to the sample for the 2006 survey round except for the addition of the new spouses of the existing respondents.

Kohler, Taulo, Masanjala *et al* (2009) explain that for the 2008 survey, some of the parents (both mothers and fathers) of the 2004 women were added to the sample in order to study intergenerational relationships.

Anglewicz, Adams, Obare *et al* express concern that some aspects of the data quality of the MDICP surveys have become important because the MDICP now encompasses five waves of collected data. A major concern is the aging of the young women who were interviewed in the early surveys because the aging causes the non-age specific measures of the latter data sets to be weighted towards the characteristics of older women. An earlier investigation of the data quality of the MDICP surveys of 1998 and 2001 by Bignami-Van Assche, Reniers and Weinreb (2003) concluded that there is a difference between people who were lost to attrition and those who remained to be interviewed in the latter survey rounds. They explain that the respondents who were lost to attrition in both surveys were more educated, had fewer children, and they were more likely to live in households with males who received an income compared to those who were not lost to attrition. In addition, Anglewicz, Adams, Obare *et al* found that in the 1998 and 2004 MDICP surveys, the respondents who were lost to attrition had fewer children and were less likely to be members of indigenous churches compared to the respondents who remained to be interviewed. Therefore, these concerns were motivation for use to investigate the fertility of the women who were interviewed in the MDICP surveys.

The article by Anglewicz, Adams, Obare *et al* (2009) presents a comparison of the mean parities of the MDICP and MDHS data sets. However, it does not adequately compare the background characteristics of the women who were interviewed in these surveys. Our investigation determines how the MDICP mean parities are affected by these characteristics, and how they compare with the mean parities for the rural women who were interviewed in the MDHSs. Two MDHSs and five MDICP surveys were held from 1998 to 2008. Hence, we will compare the 2000 MDHS data to the 1998 and 2001 MDICP data sets, and the 2004 MDHS data to the 2004, 2006 and 2008 MDICP data sets. In the next chapter, we investigate the women who were interviewed in the MDHS and MDICP surveys by the determinants of fertility.

3 DATA ANALYSIS

In this chapter, we cross-examine the data of the Malawi Diffusion and Ideational Change Project (MDICP) surveys. We shall check the consistency of the distributions of the interviewed women by four variables: age, marital status, contraceptive use, and education. We then compare these distributions with the data collected in the Malawi Demographic and Health Surveys (MDHSs). The MDHS data are used for the comparison because they are the most accurate and easily accessible source of fertility data. The investigation will indicate how and why the data sets differ. In turn, this will allow us to determine the reasons for the differences in the estimated fertility levels.

3.1 Investigation of MDICP data

As stated earlier, the comparisons of the fertility of the women who were interviewed in the MDHSs and the MDICP surveys were previously done by Anglewicz, Adams, Obare *et al* (2009). Their comparisons reveal that the mean parities for women in the MDICP surveys of 1998 and 2004 are generally higher than those shown by the 2000 and 2004 MDHSs. Their comparisons did not control for age or indeed the different age structures of the MDICP surveys. This may be one explanation for the difference in observed mean parities. Table 3.1 shows the comparison of the fertility characteristics of women who were interviewed in the MDHSs and the MDICP surveys. The mean parities for the women who were interviewed in the MDHSs are similar, while the parities for the MDICP surveys differ considerably.

Table 3.1 Mean parity and percentages for Malawi rural women, by family planning characteristics, 2000 MDHS, 2004 MDHS, 1998 MDICP, and 2004 MDICP

Characteristic	*2000 MDHS*	*1998 MDICP*	*2004 MDHS*	*2004 MDICP*
Mean number of children ever-born	3.8	4.2	3.7	4.6
Mean number of living children	2.9	3.3	3.0	3.8
Mean ideal number of children	4.5	4.8	-	-
[a]Ever use of family planning	50.0	50.4	57.2	52.5
[a]Currently using any family planning*	53.3	58.2	49.3	47.8
[a]Talked with partner about family planning	69.5	51.5	-	-
[a]Heard of family planning at clinic	63.5	88.9	-	-
[a]Heard of family planning on radio	65.1	90.1	-	-

Source: Anglewicz, Adams, Obare *et al*

Note: [a] Calculated for women aged 15-49 years. *among those who have ever used contraceptives. MDHS parities and percentages are for the rural (regional) sample. MDICP data are for the combined samples of Rumphi (northern region), Mchinji (central region), and Balaka (southern region) districts

The table also shows that the percentages of women who had ever used contraceptives in the 2000 MDHS and the 1998 MDICP data sets are similar. However, in 2004, a larger percentage of the women who were interviewed in the MDHS reported that they had ever used contraceptives compared to those who were interviewed in the MDICP survey. A larger percentage of women in the 1998 MDICP had ever heard about contraceptive methods from a clinic or on radio compared to those in the 2000 MDHS. In sum, the table shows that the mean parities for women in the 1998 and 2004 MDICP surveys were generally larger than those shown by the 2000 and 2004 MDHS data, but the comparisons of family planning characteristics do not fully substantiate the differentials. As a result, we are prompted to investigate the likely causes.

3.1.1 Characteristics of the women who were interviewed in the MDICP surveys

In this section, we examine the consistency of the recording of four background characteristics in the five survey rounds. The distributions of the women who were interviewed in the MDICP surveys by these characteristics are shown in Table 3.2.

The 1998 and 2001 surveys only recorded each woman's year of birth. Hence, each woman's age is obtained from the difference between the year of the survey and the year of birth. The other surveys recorded the age and not the year of birth. The ages of the women who were interviewed in 1998 range from 16-78 years. We obtain age ranges of 15-79 years and 0-90 years for the 2001 and 2006 surveys respectively. The 2006 age range of 0-90 years is caused by 84 women who have ages coded as 0 years, whilst the majority of the women who were interviewed during that survey round are older than 15 years. Most of the 84 women have more reasonable ages which are coded in the 2008 data set. Therefore, a more realistic age range for the women who were interviewed in 2006 is 15-82 years.

As mentioned in chapter two, the 2004 sample included never-married women for the first time and the 2008 sample included the mothers of some of the 2004 women. Hence, the 1998 and 2001 age distributions are expected to have large percentages of middle-aged women compared to the latter surveys. The 2004 distribution has a large number of women in the 15-19 age group, while the 2006 and 2008 distributions have large numbers in the oldest age groups. Our investigation also reveals that the 2004 MDICP has the largest percentage of women with missing age data.

Although the 1998 and 2001 surveys were designed to only interview ever-married women, there are nine women who were interviewed in 2001 who have missing marital status data. Consequently, we shall assume that all the women who were interviewed in this survey have in fact been married.

Table 3.2 Distributions of women by age, marital status, contraceptive use, and education, 1998 MDICP, 2001 MDICP, 2004 MDICP, 2006 MDICP, and 2008 MDICP (Percentage distributions are given in brackets)

Characteristic	*m1998 MDICP*	*m2001 MDICP*	*2004 MDICP*	*2006 MDICP*	*2008 MDICP*
Age distribution					
Age range	16-78	15-79	14-77	0-90	15-94
0-14	-	-	3(0)	88(5)	-
15-19	112(7)	26(2)	245(14)	190(10)	130(6)
20-24	300(20)	244(16)	189(11)	210(11)	333(15)
25-29	276(18)	288(18)	134(8)	256(13)	276(12)
30-34	186(12)	221(14)	116(7)	240(12)	264(12)
35-39	185(12)	226(14)	88(5)	204(11)	248(11)
40-44	109(7)	163(10)	53(3)	214(11)	216(10)
45-49	73(5)	114(7)	35(2)	141(7)	161(7)
50+	15(1)	81(5)	51(3)	218(11)	613(27)
Missing ages	280(18)	207(13)	782(46)	175(9)	0(0)
Number interviewed	1536(100)	1570(100)	1696(100)	1936(100)	2241(100)
Number not interviewed	254	383	864	936	1490
Sample size	1790	1953	2560	2872	3731
Marital status					
Ever-married	1536(100)	1561(99)	1455(86)	1784(92)	1542(69)
Never-married	-	0(0)	231(14)	123(6)	68(3)
Missing status	-	9(1)	10(1)	29(1)	631(28)
Contraceptives					
Ever used	772(50)	936(60)	718(42)	920(48)	N/A
Never used	762(50)	628(40)	890(52)	888(46)	N/A
Don't know	-	2(0)	1(0)	2(0)	N/A
Missing data	2(0)	4(0)	87(5)	126(7)	N/A
Education					
None	548(36)	521(33)	393(23)	574(30)	727(32)
Primary	905(59)	933(59)	1074(63)	1165(60)	1329(59)
Secondary	83(5)	97(6)	144(8)	163(8)	177(8)
Higher	0(0)	-	2(0)	5(0)	4(0)
Other	-	6(0)	2(0)	0(0)	-
Missing	0(0)	13(1)	81(5)	29(1)	4(0)

Source: 1998 MDICP, 2001 MDICP, 2004 MDICP, 2006 MDICP, and 2008 MDICP
Note: mEver-married women only. N/A refers to data that was not collected for a particular characteristic.

Marital status data for the latter surveys shows that the largest number of never-married women is in the 2004 survey round and that the 2008 survey has the largest number of women with missing marital status data. The large number of women with missing marital status data in the latter surveys may have been caused by the deliberate omission of the record by the interviewers, since Anglewicz, Adams, Obare *et al* assert that the interviewers relied on the data recorded in the previous survey rounds in order to locate the women who were to be interviewed in each survey round.

The distributions of women who had ever used contraceptives in the MDICP surveys of 1998, 2001, 2004, and 2006 are derived from a question on whether each woman had ever used any traditional or modern method for family planning or birth spacing. During the 2008 survey round, no question was asked about contraceptive use. Therefore, the distribution is not shown in Table 3.2. The available data shows that the majority of the women who were interviewed in the 2001 MDICP had ever used contraceptives. The distributions for the latter surveys reveal no clear pattern. Thus, it is difficult to study the pattern of contraceptive use over time because the study is hugely affected by the large percentages of women with missing data.

The distributions of the level of education attained by the women in the 1998 and 2001 surveys are obtained from the responses to two questions. The first question determines whether a woman had ever been to school and the second question asks for the highest level of education she attained. The total number of women with missing education data is then derived from the difference between the number of women in total and the combined women who responded to the two questions. The 2004 data set categorised the women with no education and those with only primary education in the same category and a second variable records the number of years at the recorded level of education. Therefore, women with no education are coded as “none or only primary education” in the first variable and“0-level”of education in the second variable.

Our major concern is the large percentage of women in the 2004 survey whose ages were not recorded. Leaving them as such gives a small number of women in the oldest and youngest age groups and this might cause biased fertility estimates. In addition, we also note that the 1998 and 2001 data sets consist of only ever-married women. Hence, the fertility estimates of these two data sets are likely to be higher relative to the estimates for the other data sets because ever-married women generally have higher fertility than never-married women. Our investigation might be hindered by the unrecorded contraceptive data in the 2008 survey and the poorly categorised education data in the 2004 survey. Our conclusion is that the recording of the characteristics which we investigated is not consistent throughout the five surveys. Therefore, we cannot conduct an investigation of the fertility levels implied by the MDICP data without correcting these differences.

3.1.2 Age

We shall estimate the missing ages of the women who were interviewed in each survey from their recorded ages in the other survey rounds. We consider each woman’s age to be consistently recorded in two surveys if the difference between the ages is equal to the time between the two surveys. Therefore, the proportion of consistent ages in the two surveys will determine the accuracy of imputing the missing ages across the two surveys.

Table 3.3 Proportions of consistent ages before imputing missing ages, 1998 MDICP, 2001 MDICP, 2004 MDICP, 2006 MDICP, and 2008 MDICP

	Second survey			
First survey	*2001*	*2004*	*2006*	*2008*
1998	0.68	0.31	0.59	0.59
2001		0.33	0.59	0.49
2004			0.99	1.00
2006				0.97

Source: 1998 MDICP, 2001 MDICP, 2004 MDICP, 2006 MDICP, and 2008 MDICP

The proportion of women with consistent ages in each pair of MDICP surveys is given in each cell of Table 3.3. The table shows that there are larger proportions of consistent ages for the paired latter surveys relative to the paired early survey rounds. Table 3.4 shows the distributions of the differences between the observed and estimated ages of the women who were interviewed in consecutive MDICP surveys.

Table 3.4 Distributions of differences between estimated and observed ages, by period (Percentage distributions are given in brackets)

[a]Difference	*Period*			
	1998-2001	*2001-2004*	*2004-2006*	*2006-2008*
<-5	27(3)	25(5)	1(0)	82(0)
-5	4(0)	4(1)	0(0)	0(0)
-4	10(1)	5(1)	0(0)	0(0)
-3	12(1)	9(2)	1(0)	0(0)
-2	32(3)	18(3)	2(0)	0(0)
-1	64(7)	42(8)	0(0)	0(0)
0	637(68)	170(33)	1014(99)	2439(97)
1	71(8)	61(12)	0(0)	0(0)
2	31(3)	40(8)	1(0)	0(0)
3	18(2)	33(6)	0(0)	0(0)
4	7(1)	25(5)	0(0)	0(0)
5	9(1)	12(2)	0(0)	0(0)
>5	18(2)	72(14)	1(0)	0(0)
Consistent	637(68)	170(33)	1014(99)	2439(97)
Inconsistent	303(32)	346(67)	6(1)	82(3)
Total	940(100)	516(100)	1020(100)	2521(100)

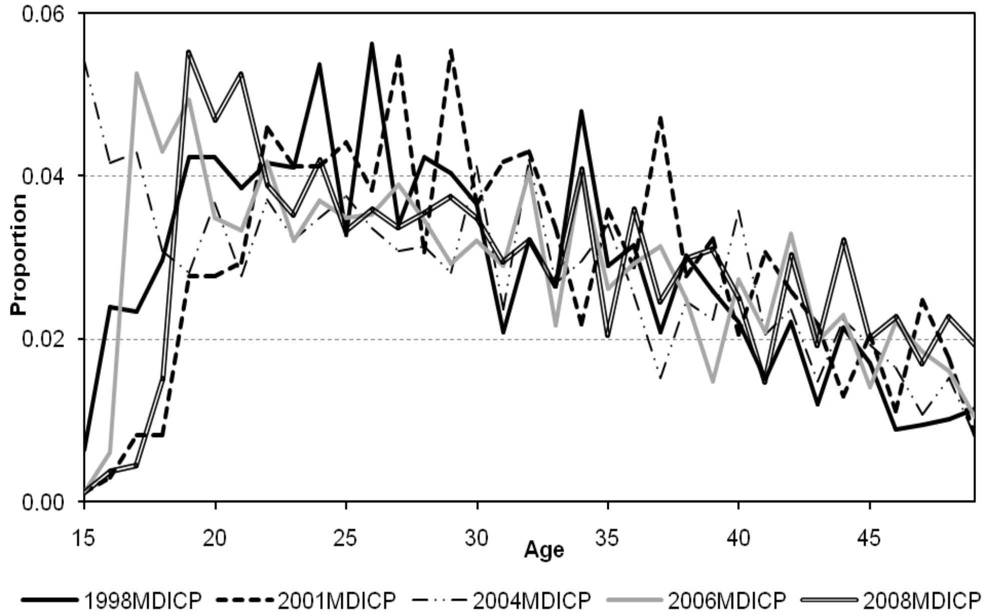
Source: 1998 MDICP, 2001 MDICP, 2004 MDICP, 2006 MDICP, and 2008 MDICP

Note: The estimated age of each woman in the first survey is calculated by subtracting the length of the inter-survey period from the recorded age in the second survey. [a] calculated by subtracting the estimated age from the recorded age in the first of two surveys. Each woman's age is considered to be consistently recorded in two surveys if the difference between her estimated age and her observed age is zero.

The estimated age of each woman in the first survey is calculated by subtracting the length of the inter-survey period from the recorded age in the second survey. Very few women have recorded ages in both surveys of the two early inter-survey periods. The distribution of the 1998-2001 surveys is aproximately symetric around zero and this suggests that the error of inconsistent ages is random for the two surveys. The table shows that the 2001 and 2004 surveys have larger percentages of women with positive differences and this shows that the inconsistent ages in the 2001 survey were generally larger than the estimated ages obtained from the 2004 surveys.

Since there are very few missing ages in the 2006 and 2008 surveys, we shall first impute the missing ages in the 2006 and 2008 surveys and then we use them to estimate the missing ages in each of the previous surveys. We conclude the imputing by editing all the ages to ensure that they are consistent across all the five survey rounds.

Figure 3.1 Age distributions of women aged 15-49, who were interviewed in the MDICP surveys, 1998 MDICP, 2001 MDICP, 2004 MDICP, 2006 MDICP, and 2008 MDICP



Source: 1998 MDICP, 2001 MDICP, 2004 MDICP, 2006 MDICP, and 2008 MDICP
Note: 1998 and 2001 MDICP distributions are for women who have ever been married.

The resultant age distributions of the women aged 15-49 years who were interviewed in the five MDICP surveys are shown in Figure 3.1. The effect of adding the extra adolescents and unmarried women is clearly evident in the age distribution for the 2004 round as is the addition of the mothers of some younger women in 2008 who themselves may be still of reproductive age. There is a preference for even numbered ages by older women in all the data sets.

3.1.3 Marriage

In order to ensure that the marital status data of each woman is consistently recorded in all the five data sets, we recoded the data to ensure that every woman who was interviewed in the 1998 and 2001 surveys is recorded to have ever been married in all the five survey rounds. This is necessary since we know that all the women of these two surveys have ever been married. We also ensured that every woman is recorded to have ever-married in all the latter survey rounds if she is recorded to have ever-married in an earlier round. As ever being married is strongly correlated with age, it is not surprising that the smallest proportion of ever-married women is at the two youngest ages as shown in Table 3.5.

Table 3.5 Proportions of ever-married rural women by age group and survey, 2004 MDICP, 2006 MDICP, and 2008 MDICP

Age group	*2004 MDICP*	*2006 MDICP*	*2008 MDICP*
15-19	0.29	0.60	0.48
20-24	0.93	0.95	0.81
25-29	1.00	0.99	0.98
30-34	1.00	1.00	0.99
35-39	1.00	1.00	1.00
40-44	1.00	1.00	1.00
45-49	1.00	1.00	1.00

Source: 2004 MDICP, 2006 MDICP, and 2008 MDICP

Note: 1998 MDICP and 2001 MDICP surveys are not included because they interviewed ever-married women only.

Despite this, we observe that the 2006 survey has a relatively larger proportion of ever-married women who are aged 15-19 years. This is probably caused by new marriages to some of the adolescent girls who were added to the study in 2004. The proportion of ever-married women in the youngest age group of the 2008 MDICP is smaller than expected and this is probably caused by the fact that the there are very few women in the age group relative to the number for the other data sets.

3.1.4 Marriage and contraceptive use

In Table 3.6, the overall proportion of ever-married women who ever used contraceptives (either traditional or modern methods) peaks in 2006 and the age specific proportions for older women in 2006 are higher relative to the corresponding proportions for the other surveys. We also observe that the proportions for women who are older than 40 years in 2004 and 2006 are generally higher than the corresponding proportions in both the previous surveys. The most obvious explanation is that those women who reported to have previously never used contraceptives when they were younger in 1998 and 2001 report that they used contraceptives in these two latter surveys. We repeat the investigation by using the data of women who were interviewed in all the previous survey rounds.

Table 3.6 Proportions of ever-married women who ever used modern or traditional contraceptives, by age group and survey

Age group	*1998 MDICP*	*2001 MDICP*	*2004 MDICP*	*2006 MDICP*
15-19	0.23	0.53	0.21	0.32
20-24	0.48	0.66	0.66	0.66
25-29	0.56	0.77	0.80	0.87
30-34	0.58	0.78	0.87	0.87
35-39	0.62	0.72	0.88	0.91
40-44	0.54	0.76	0.80	0.82
45-49	0.43	0.74	0.79	0.82
Overall	0.50	0.73	0.77	0.79

Source: 1998 MDICP, 2001 MDICP, 2004 MDICP and 2006 MDICP
Note: 2008 MDICP did not collect data on contraceptives use.

The proportions of ever-married women who ever used contraceptives and were interviewed in all prior surveys are shown in Table 3.7. The proportions for women who were older than 40 years in 2006 are slightly higher compared to corresponding proportions in the early surveys, while the proportions for young women increase substantially when the data are restricted to women who were interviewed in all surveys.

Table 3.7 Proportions of ever-married women aged 15-49 who ever used modern or traditional contraceptives and were interviewed in all previous survey rounds, by age group and survey

Age group	*1998 MDICP*	*2001 MDICP*	*2004 MDICP*	*2006 MDICP*
15-19	0.23	0.59		
20-24	0.48	0.70	0.79	0.81
25-29	0.56	0.81	0.83	0.94
30-34	0.58	0.77	0.88	0.90
35-39	0.62	0.75	0.88	0.92
40-44	0.54	0.78	0.81	0.83
45-49	0.43	0.75	0.80	0.85
Total interviewed	1463(100*)	1340(92*)	1233(84*)	1183(81*)

Source: 1998 MDICP, 2001 MDICP, 2004 MDICP, and 2006 MDICP
Note: The 2008 MDICP survey did not collect data on contraceptives use. *Refers to percentage of the original number of women who were interviewed in 1998. Proportion for the 15-19 age group is not shown for the 2004 and 2006 surveys because all of the women who were in that age group in previous surveys had grown older.

Our observations show that the data are now consistent because we expect the young respondents who report that they never used contraceptives during the early surveys to report that they had used contraceptives in the latter surveys, while we do not expect the proportions for older women to change during the study period.

3.1.5 Education

The proportions of ever-married women with some education are shown in Table 3.8. The overall proportions are larger during the latter years compared to the 1998 and 2001 proportions. As expected, the proportions for 2004, 2006 and 2008 are higher relative to the proportions for the previous surveys because of the adolescent girls (who most probably had some education) who were added in 2004. We proceed to investigate by using the data of women who were interviewed in all the previous survey rounds.

Table 3.8 Proportions of ever-married women with some education, by age group and survey

Age group	1998 MDICP	2001 MDICP	2004 MDICP	2006 MDICP	2008 MDICP
15-19	0.78	0.83	0.97	0.88	0.97
20-24	0.75	0.85	0.88	0.92	0.95
25-29	0.67	0.78	0.88	0.90	0.93
30-34	0.62	0.72	0.81	0.83	0.89
35-39	0.58	0.67	0.80	0.79	0.81
40-44	0.49	0.64	0.73	0.77	0.77
45-49	0.54	0.66	0.76	0.75	0.78
Overall	0.66	0.74	0.89	0.84	0.87

Source: 1998 MDICP, 2001 MDICP, 2006 MDICP, and 2008 MDICP

The proportions of ever-married women with some education, who were interviewed in all prior survey rounds, are shown in Table 3.9. The total number of women decreases substantially between 1998 and 2008; only 77 per cent of the original number of women have available data in 2008.

Table 3.9 Proportions of ever-married women with some education, who were interviewed in all previous survey rounds, by age group and survey

Age group	1998 MDICP	2001 MDICP	2004 MDICP	2006 MDICP	2008 MDICP
15-19	0.78	0.81			
20-24	0.75	0.85	0.85	0.90	
25-29	0.67	0.80	0.88	0.89	0.90
30-34	0.62	0.74	0.83	0.85	0.90
35-39	0.58	0.67	0.79	0.80	0.83
40-44	0.49	0.64	0.72	0.78	0.81
45-49	0.54	0.67	0.75	0.75	0.78
Total interviewed	1466(100*)	1408(88*)	1283(88*)	1235(84*)	1126(77*)

Source: 1998 MDICP, 2001 MDICP, 2004 MDICP, 2006 MDICP, and 2008 MDICP
Note: *Refers to the percentage of women who were interviewed relative to the number of women who were interviewed in 1998. The missing proportions for the 15-19 age group are omitted because most of the respondents in the age group had grown older.

In Table 3.9, we also observe that the age specific proportions of women with some education, who were interviewed in all prior surveys and were aged from 20-44 years in 2001 increase substantially to the proportions aged 25-49 years in 2006.We do not expect the proportions to increase substantially over time especially for older women. Hence the substantial increase is most likely to have been caused by our data imputations for 2004.

3.2 Data from the MDHS, and comparison with data from MDICP

Demographic and Health Surveys interview both never-married and ever-married women who are aged 15-49 years. Unlike the MDICP surveys, the women who were interviewed in the MDHSs were sampled from both urban and rural areas. In addition, MDHS data are weighted to make them nationally representative. The overall proportion of urban women increased slightly from the proportion which was interviewed in the 2000 MDHS to the proportion interviewed in the 2004 MDHS. Table 3.10 shows the proportion of urban women who were interviewed in the MDHSs.

Table 3.10 Proportions of women living in urban areas by age group and survey, 2000 MDHS and 2004 MDHS

Age group	*2000 MDHS*	*2004 MDHS*
15-19	0.17	0.19
20-24	0.19	0.21
25-29	0.16	0.19
30-34	0.16	0.14
35-39	0.13	0.16
40-44	0.12	0.14
45-49	0.11	0.14
All ages	0.16	0.18

Source: 2000 MDHS and 2004 MDHS

The exclusion of women who live in urban areas is important since Harwood-Lejeune (2000) explains that, in sub-Saharan Africa, women who live in urban areas generally have lower fertility than women who live in rural areas. Hence, our comparisons of the MDHS and MDICP data are restricted to rural women for the MDHS data. Likewise, in drawing comparisons with the MDICP data (especially the earlier rounds), the MDHS needs to be analysed for ever-married rural women only. In order to determine the causes of the fertility differential, observed by Anglewicz, Adams, Obare *et al*, we shall use regional level data from the MDHS data sets.

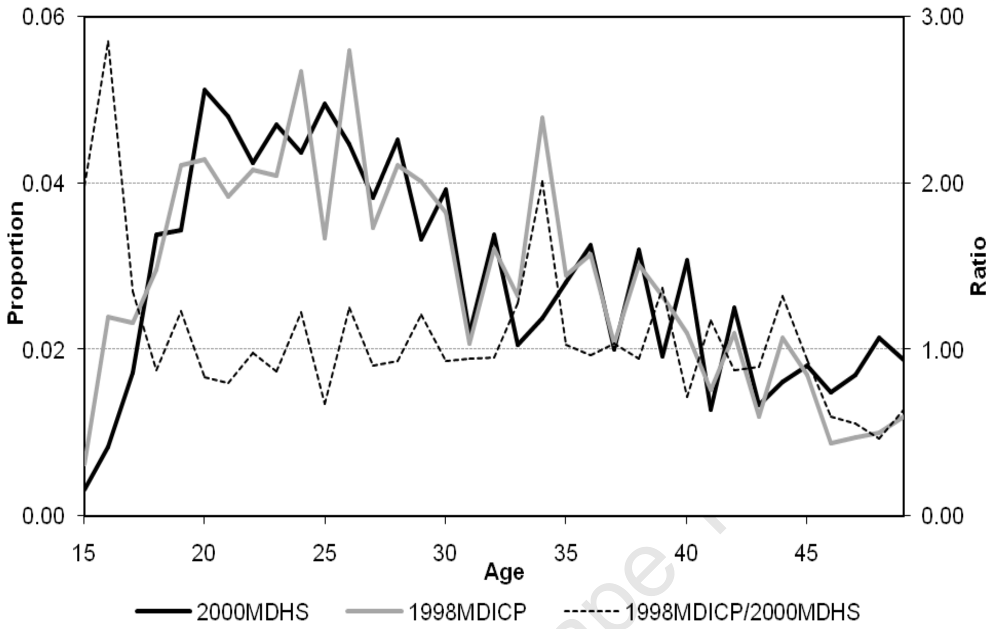
3.2.1 Age distributions in the MDHS and MDICP data

To investigate whether the age distributions of the women who were interviewed in the MDHSs and the MDICP surveys are confounded by age, we compare the distributions graphically by using the data of women who are aged 15-49 years. Our comparisons are hindered by the availability of the data collected from only two MDHSs. Therefore, we compare the 2000 MDHS distribution to the 1998 and 2001 MDICP distributions, and the 2004 MDHS distribution to the 2004, 2006, and 2008 MDICP age distributions.

The age distributions of ever-married rural women who are aged 15-49 years, who were interviewed in the 2000 MDHS and 1998 MDICP surveys are shown in Figure 3.2.

The line showing the ratio of the two distributions fluctuates around one. Hence, the two distributions are very similar.

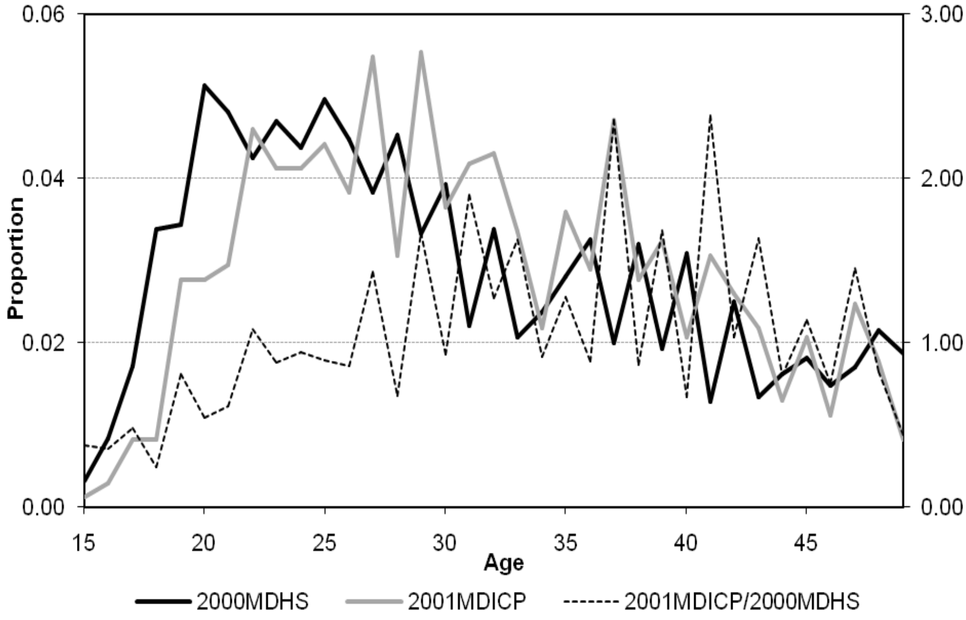
Figure 3.2 Age distributions of ever-married rural women aged 15-49, 2000 MDHS and 1998 MDICP



Source: 1998 MDICP and 2000 MDHS

In Figure 3.3, we compare the age distribution of ever-married rural women in the 2001 MDICP survey to the distribution for ever-married rural women in the 2000 MDHS. The MDHS distribution has larger proportions of women aged 15-20 years, while the MDICP has larger proportions of older women. The women who were interviwed in the 2001 survey are expected to be older than 19 years because most of the women who were interviwed in 1998 were older than 16 years.

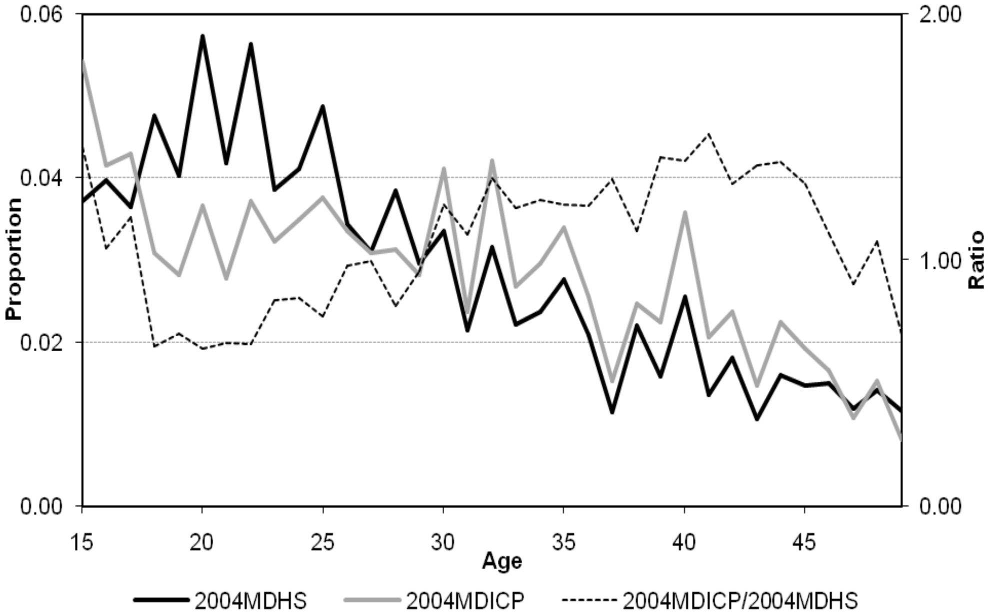
Figure 3.3 Age distributions of ever-married rural women aged 15-49, 2000 MDHS and 2001 MDICP



Source: 2000 MDHS and 2001 MDICP

Our comparison of the 2004 MDHS and the 2004 MDICP age distributions of rural women in Figure 3.4 shows that there are larger proportions of very young and middle-aged women in the MDICP survey relative to the MDHS. We also observe that the MDHS proportion of women aged 16-25 years is larger relative to the MDICP proportion. The ratio of the two distributions is generally greater than one for women who are older than 24 years and it is less than one for younger women. There are larger proportions of women in the youngest age group of the MDICP age distribution because approximately 90 per cent of the girls who were added to the study in 2004 were aged between 12 and 20 years. We also observe that the general shapes of the two distributions are similar at the oldest ages. However, the MDICP proportions are marginally larger becasuse of the aging of the original sample of the women who were interviwed in the 1998 survey.

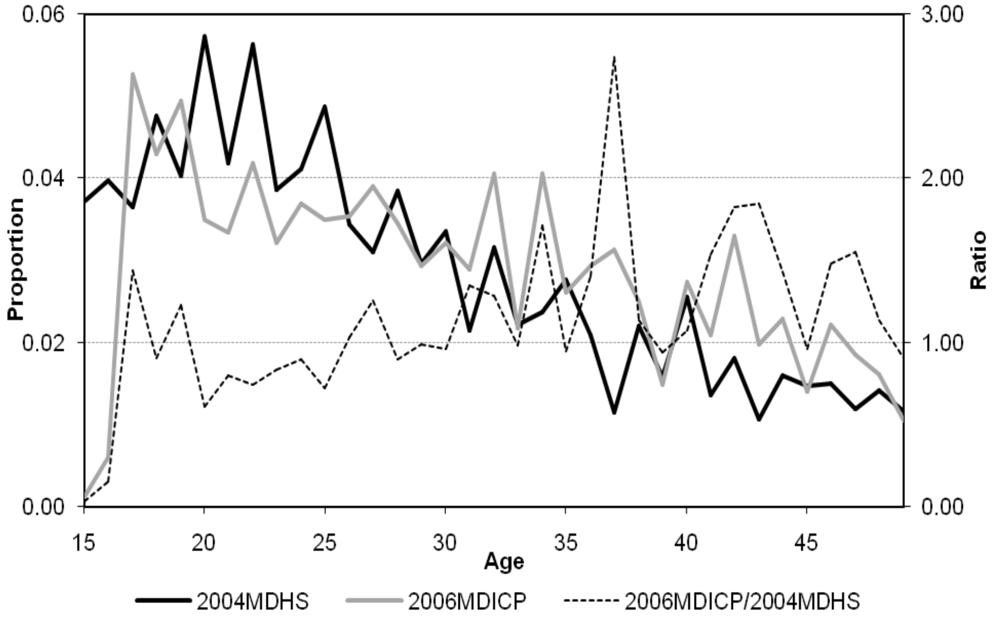
Figure 3.4 Age distributions of rural women aged 15-49, 2004 MDHS and 2004 MDICP



Source: 2004 MDICP and 2004 MDHS

A comparison of the age distributions of the rural women who were interviwed in the 2004 MDHS and the 2006 MDICP survey is shown in Figure 3.5. The MDHS distribution shows larger proportions of women who are younger than 20 years, whilst the MDICP distribution has larger proportions of women who are older than 30 years. If we compare the 2004 MDICP age distribution in Figure 3.4 to the 2006 MDICP distribution in Figure 3.5, we observe that the two distributions are similar except that the 2006 structure moved two years to the right. This shift is caused by the gradual aging of the women who were interviewed in the 2004 survey round. Hence, this results in a larger proportion of older women who are shown by the MDICP distribution relative to the MDHS distribution as shown in Figure 3.5.

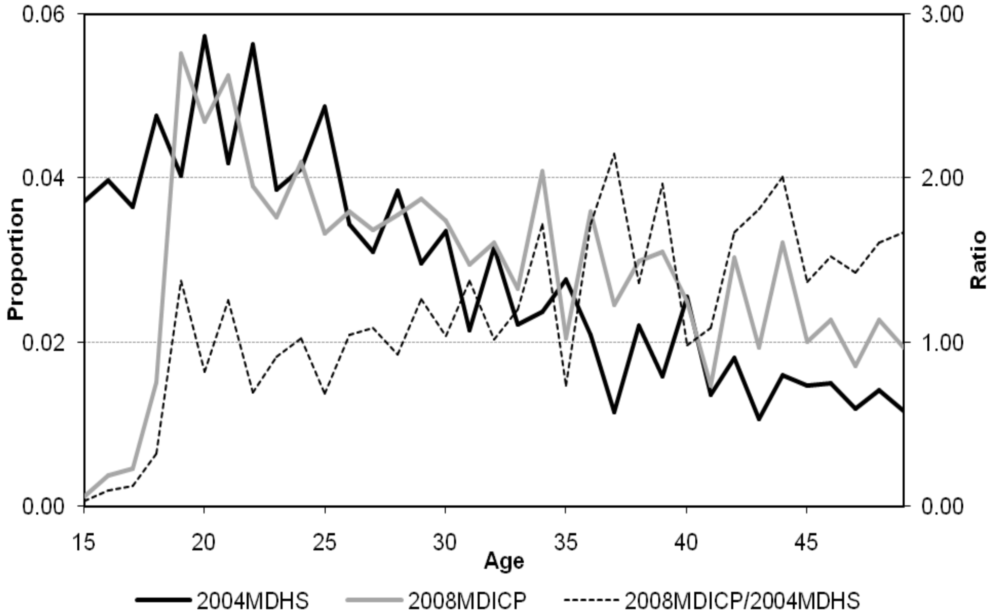
Figure 3.5 Age distributions of all the rural women aged 15-49, 2004 MDHS and 2006 MDICP



Source: 2004 MDHS and 2006 MDICP

In Figure 3.6, the ratio of the age distributions of the women who were interviwed in the 2008 MDICP to those in the 2004 MDHS is generally greater than one at the oldest ages. This shows that the MDHS distribution has a large proportion of young women and a small proportion of older women relative to the MDICP distribution.

Figure 3.6 Age distributions of all the rural women aged 15-49, 2004 MDHS and 2008 MDICP



Source: 2004 MDHS and 2008 MDICP

Both distributions in Figure 3.6 tail off at the oldest ages but the MDHS distribution tails off at a faster rate. The MDICP distribution fluctuates more visibly for women who are older than 34 years and it has larger proportions of women in this age range. It is likely that the large proportions of older women are caused by the addition of mothers to the study or the aging of the women from the previous survey rounds.

In sum, only the 2000 MDHS and the 1998 MDICP age distributions show no major difference. Therefore, we expect the effect of age confounding on the fertility comparisons between the 1998 MDICP and 2000 MDHS data to be negligible. We also note that the latter MDICP surveys interviwed larger proportions of older women relative to the MDHS because of the gradual aging of the women who were interviwed in the early survey rounds. Therefore, we expect the women who were interviewed in the latter MDICP surveys to have higher overall mean parities than the women who were interviewed in the latter MDHSs because older women generally have higher parities than younger women.

3.2.2 Proportions of ever-married rural women in the MDHS and MDICP data

The proportions of ever-married rural women who were interviewed in the MDHSs and the MDICP surveys are shown in Table 3.11. Since the 1998 and 2001 MDICP surveys only interviwed ever-married women, they are not included in the table.

Table 3.11 Proportions of ever-married rural women by age group and survey

Age group	*2000 MDHS*	*2004 MDHS*	*2004 MDICP*	*2006 MDICP*	*2008 MDICP*
15-19	0.38	0.39	**0.29	**0.60	**0.48
20-24	0.91	0.92	0.93	*0.95	0.81
25-29	0.99	0.98	**1.00	**0.99	**0.98
30-34	0.99	0.99	**1.00	1.00	**0.99
35-39	1.00	0.99	**1.00	**1.00	**1.00
40-44	0.99	1.00	**1.00	**1.00	**1.00
45-49	1.00	0.99	**1.00	**1.00	**1.00

Source: 2000 MDHS, 2004 MDHS, 2004 MDICP, 2006 MDICP, and 2008 MDICP
Note: 1998 and 2001 MDICP are not included because they interviewed ever-married women only. 2000 MDHS is compared to 2004 MDICP, while the 2004 MDHS is compared to the 2006 and the 2008 MDICP. Chi-squared test for equality of proportions was conducted: *means significantly different at the 5% level of significance. ** Significantly different at the 1% level of significance

In the table, we compare the age specific proportions of the 2000 MDHS to the 2004 MDICP proportins and the 2004 MDHS to the 2006 and 2008 MDICP proportions. The proportions of ever-married women aged 15-19 years are significantly higher for the two MDHSs compared to the proportions for women in the 2004 MDICP. This is probably caused by the unbalanced addition of never-married adolescents to the 2004 MDICP. The age specific proportions for the five surveys show that the majority of the women who are aged 25-49 years have ever been married. Therefore, based on this evidence, we expect the highest fertility to be observed for women who are older than 25 years in both surveys.

3.2.3 Proportions of ever-married women who ever used contraceptives in the MDHS and MDICP data

The age specific proportions of ever-married women who have ever used contraceptives in the MDHSs and the MDICP surveys are shown in Table 3.12. We compare the proportions for the two early MDICP surveys to the proportions for the early MDHS and the proportions for the latter MDICP surveys to the 2004 MDHS. The overall proportion for the 2001 MDICP data set is higher relative to the proportion for the 2000 MDHS, while the proportions for the 2000 MDHS and the 1998 MDICP, and those for the 2004 MDHS and the 2006 MDICP data sets are not significantly different from each other.

Table 3.12 Proportions of ever-married rural women who have ever used contraceptives, by age group and survey

Age group	*2000 MDHS*	*2004 MDHS*	*1998 MDICP*	*2001 MDICP*	*2004 MDICP*	*2006 MDICP*
15-19	0.27	0.33	0.23	*0.53	*0.21	**0.32
20-24	0.46	0.54	*0.48	*0.66	**0.66	**0.66
25-29	0.57	0.66	0.56	*0.77	**0.80	**0.87
30-34	0.58	0.66	0.58	*0.78	**0.87	**0.87
35-39	0.54	0.65	0.62	*0.72	**0.88	**0.91
40-44	0.54	0.63	0.54	*0.76	**0.80	**0.82
45-49	0.45	0.48	0.43	*0.74	**0.79	**0.82
Overall	0.50	0.58	*0.50	*0.73	**0.77	**0.79

Source: 2000 MDHS, 2004 MDHS, 1998 MDICP, 2001 MDICP, 2004 MDICP, and 2006 MDICP

Note: All the proportions are for rural women who have ever been married. The 2008 MDICP is not included because it did not collect data on contraceptive use. The 2000 MDHS proportions are compared to the 1998 and 2001 MDICP proportions and the 2004 MDHS proportions are compared to the 2004 and 2006 MDICP proportions. Chi-squared test conducted: * means significantly different at the 5% level of significance. **Significantly different at the 1% level of significance.

Generally, the table shows that the proportion of women who had ever used contraceptives increased over time and we expect this increase to favour lower fertility estimates for the latter surveys. We expect the effect of contraceptive use on the fertility of women in the 2000 MDHS and the 1998 MDICP data sets to be similar since both data sets have comparable proportions of women who have ever used contraceptives. The table confirms that the proportions of women in the MDICP who had ever used contraceptives increased at a faster rate relative to the MDHS proportions. Thus, we expect the fertility of the women in the latter surveys to be lower.

3.2.4 Proportions of women with some education in the MDHS and MDICP data

The overall proportion of women with some education in the 2000 MDHS data set is similar to the proportion for the 1998 MDICP, while the 2006 and 2008 proportions are larger than the MDHS proportions. Our comparison of the 2000 MDHS to the 1998 and 2001 MDICP shows that the age specific proportions of educated women in the 2001 MDICP are generally higher than the proportions for the 2000 MDHS. We also observe that the age specific proportions are significantly higher for the 2004, 2006 and 2008 MDICP relative to the proportions for the 2004 MDHS.

Table 3.13 Proportions of women aged 15-49 who have some education, by age group and survey

Age group	*m2000 MDHS*	*2004 MDHS*	*m1998 MDICP*	*m2001 MDICP*	*2004 MDICP*	*2006 MDICP*	*2008 MDICP*
15-19	0.84	0.94	*0.78	*0.83	*0.97	0.95	**0.99
20-24	0.78	0.86	0.75	**0.85	*0.89	**0.93	**0.97
25-29	0.66	0.73	**0.67	**0.78	**0.88	**0.90	**0.93
30-34	0.61	0.61	0.62	**0.72	**0.81	**0.83	**0.89
35-39	0.56	0.58	**0.58	**0.67	**0.80	**0.79	**0.81
40-44	0.50	0.58	**0.49	**0.64	**0.73	**0.77	**0.78
45-49	0.47	0.47	0.54	0.66	**0.76	**0.75	**0.78
Overall	0.65	0.74	**0.66	**0.74	**0.85	**0.86	**0.88

Source: 2000MDHS, 2004MDHS, 1998MDICP, 2001MDICP, 2004MDICP, 2006MDICP, and 2008MDICP.

Note: mEver-married rural women only. Chi-squared test conducted: *means significantly different from the corresponding MDHS proportion at the 5% level of significance. ** significantly different from the corresponding MDHS proportion at the 1% level of significance

The pattern for both surveys indicates that the overall proportion of women with some education increased over the period from 1998 to 2008. Therefore, we expect the fertility trend of both surveys to show a declining pattern. In Table 3.13, we compare the proportions of rural women with some education in the MDHSs to the proportions for women in the MDICP surveys. We proceed to interrogate the data of ever-married rural women only.

Table 3.14 Proportions of ever-married rural women aged 15-49 who have some education, by age group and survey

Age group	*2000 MDHS*	*2004 MDHS*	*1998 MDICP*	*2001 MDICP*	*2004 MDICP*	*2006 MDICP*	*2008 MDICP*
15-19	0.84	0.89	*0.78	*0.83	0.97	**0.88	0.97
20-24	0.78	0.85	0.75	*0.85	0.88	**0.92	**0.95
25-29	0.66	0.73	**0.67	*0.78	**0.88	**0.90	**0.93
30-34	0.61	0.61	0.62	*0.72	**0.81	**0.83	**0.89
35-39	0.56	0.58	**0.58	*0.67	**0.80	**0.79	**0.81
40-44	0.50	0.58	**0.49	*0.64	**0.73	**0.77	**0.77
45-49	0.47	0.46	0.54	0.66	**0.76	**0.75	**0.78
Overall	0.65	0.70	**0.66	*0.74	**0.83	**0.84	**0.87

Source: 2000MDHS, 2004MDHS, 1998MDICP, 2001MDICP,2004MDICP, 2006MDICP,2008MDICP.

Note: **Significantly different from the corresponding MDHS proportion at the 1% level of significance.

In Table 3.14, we compare the proportions of ever-married rural women with some education in the MDHSs to the proportions for women in the MDICP surveys. There are no substantial differences between the resulting proportions for the MDICP data in Table 3.14 compared to the proportions shown in Table 3.13 except for the 2004 proportions. Therefore, the proportions of women with some education in the MDICP surveys are similar for the data sets of all the women (both never-married and ever-married) and the corresponding data sets of ever-married women.

3.2.5 Summary of findings and their expected implications

Since the MDICP study included never-married women for the first time in 2004 and then introduced the mothers of some of the women during the 2008 round, the MDHS data sets have to be restricted accordingly. This is done in order to allow for reasonable comparisons to be carried out between the data of the two respective surveys. As a result, we restrict the 2000 MDHS data set to ever-married rural women in order for us to be able to compare it to the 1998 and 2001 MDICP data sets. We can compare the fertility of women in the 2004 MDHS to the fertility of women in the 2004, 2006, and 2008 MDICP surveys without adjusting the data sets. However, in order to interrogate the fertility implied by the MDICP data over time, we shall use the data of ever-married rural women for all the survey rounds because we expect the number of never-married women to decline substantially between 2004 and 2008 because of new marriages to the girls who were added to the study in 2004.

Our investigations show that, after making the necessary adjustments to the MDHS and the MDICP data sets, the age distributions of the women who were interviewed in the MDICP surveys still differ from the distributions for the women who were interviewed in the MDHSs. The differences observed for the latter surveys occur as the initial sample of ever-married women ages. Thus, there were very few women who were aged 15-19 years during the 2001 survey. When never-married woman are included for the first time in 2004, the proportion of women aged 15-19 years increases and it becomes disproportionately larger than the proportions for older age groups. The distribution for the 2008 survey shows that the proportion of older women increases with the inclusion of some of the women's mothers. However, the age distributions of the 2000 and 2004 MDHSs have similar patterns which resemble the population distribution of the women in the country. Since the MDICP is a longitudinal study, we expected that the age distributions of the women who were interviewed in the latter surveys to have larger proportions of older women, and the overall mean parities of the women in these surveys to be weighted towards the fertility of older women.

Our summary of the proportions of ever-married women in Table 3.11 shows that all the women in the MDICP surveys got married when they were aged 25-29 years. We also observed that the proportions of ever-married women who ever used contraceptives are similar for the 2000 MDHS and 1998 MDICP, while the proportions are larger for the 2001 MDICP relative to the 2000 MDHS. In general, the overall proportions of women who were interviewed in both surveys, who had ever used contraceptives increased between 1998 and 2008 and this is consistent with a decline in fertility.

We expect the 1998 MDICP and the 2000 MDHS data sets to have similar age specific mean parities because the age distributions of the women who were interviewed in the two surveys are similar. However, we expect the 2001 MDICP to have larger mean parity measures than the 2000 MDHS because it has very few women in the less reproductive 15-19 age group. In addition, Figure 3.6 shows that the proportions of older women in the 2008 MDICP are larger relative to those in the 2004 MDHS. Hence the mean parities implied by the MDICP data sets are expected to rise during the period from 2004 to 2008, while the increase in the proportion of women with some education is expected reduce the magnitude of the expected fertility increase.

In sum, we expect the age distribution and the proportion of ever-married women to be the major determinants of the fertility differentials between the MDICP and MDHS data. We control for these two variables in the next section of our investigation.

3.3 Mean parities of the MDICP and MDHS data

We begin by testing the parity data of the women who were interviewed in the MDICP surveys for inter-survey consistency. Then, we estimate the mean parities for the women who were interviewed in the MDICP surveys in order to compare them with the parities for women in the MDHSs. We use the fact that parity increases monotonically between successive years. We also use the rule of thumb that each woman can only give birth to a maximum of two children in three years. Hence, all the women whose parities decrease in consecutive surveys or women whose parities increase by a rate of more than two children in three years are inconsistent. Table 3.15 presents the distributions of the women by inter-survey period and parity increment.

Table 3.15 Distributions of parity increment by MDICP inter-survey period (Percentage distributions are given in brackets)

Parity increment	*Period* 1998-2001	2001-2004	2004-2006	2006-2008
<-3	27(2)	42(4)	53(4)	27(2)
-2	33(3)	38(3)	36(3)	32(2)
-1	79(7)	87(7)	78(6)	81(6)
0	349(29)	412(35)	664(47)	666(48)
1	484(40)	423(36)	402(28)	442(32)
2	160(13)	113(9)	105(7)	89(6)
>3	80(7)	76(6)	75(5)	58(4)
Consistent	993(82)	948(80)	1066(75)	1108(79)
Inconsistent	219(18)	243(20)	347(25)	287(21)
Total	1212(100)	1191(100)	1413(100)	1395(100)

Source: 1998 MDICP, 2001 MDICP, 2004 MDICP, 2006 MDICP, and 2008 MDICP

Approximately 20 per cent of the women have inconsistent parities in the four inter-survey periods. This shows that the proportion of erroneous parity data is constant throughout the inter-survey periods. Inconsistent parities cannot be corrected because we cannot determine the survey in which the error occurred. Therefore, we exclude the data of women with inconsistent parities from our fertility estimations[1].

Table 3.16 Mean parities of ever-married rural women aged 15-49, by age group and survey

Age group	*2000 MDHS*	*2004 MDHS*	*1998 MDICP*	*2001 MDICP*	*2004 MDICP*	*2006 MDICP*	*2008 MDICP*
15-19	0.8	0.7	1.2	2.1	0.9	1.1	1.6
20-24	1.8	1.9	2.0	2.3	2.3	2.1	2.2
25-29	3.2	3.2	3.5	3.5	3.7	3.7	3.5
30-34	4.5	4.5	5.2	5.1	4.8	4.9	5.0
35-39	5.6	5.6	6.4	6.1	6.2	5.9	6.0
40-44	6.6	6.4	7.0	7.2	7.1	6.7	6.5
45-49	7.1	6.9	6.6	7.6	7.7	7.6	7.6
Overall Parity							
Unweighted	3.8	3.7	4.0	4.7	4.7	4.5	4.8
Weighted	-	-	4.2[a]	4.4[a]	4.2[b]	4.1[b]	4.1[b]

Source: 2000MDHS, 2004MDHS, 1998MDICP, 2001MDICP, 2004MDICP, 2006MDICP, 2008MDICP

Note: MDICP mean parities differ with those calculated by Anglewicz, Adams, Obare *et al* in Table 3.1 because we used consistent age and parity data according to the conditions set in Section 3.1.2 and at the beginning of this section about the general rules for consistent age and consistent party data respectively. [a]Derived by using weights from the 2000 MDHS age distribution. [b]Derived by using weights from the 2004 MDHS age distribution.

We calculate the weighted mean parity of each MDICP data set in order to remove the effect of age confounding.

[1] An alternative to dropping all the women whose reported parity is inconsistent is to use data from the household roster and family listing in 2006 and 2008, which lists all the children the women had, by year, whether living or not.

The weighted mean is calculated by using the age specific mean parities of the women who were interviewed in each MDICP survey and the weights from the age distribution of ever-married rural women who were interviewed in the corresponding MDHS data set. In Table 3.16, we compare the mean parities of ever-married rural women who have consistent age and parity data to the mean parities for the corresponding women in the 2000 and 2004 MDHSs.

We have restricted the 2000 and 2004 MDHS data sets to ever-married rural women in order to obtain mean parities which are comparable to the parities implied by the MDICP data sets. The overall mean parity (unweighted) of the MDICP surveys fluctuates during the study period. The 2001 and 2008 MDICP surveys record the highest overall mean parities (unweighted) relative to the other MDICP surveys. The lowest overall mean parity (unweighted) for the 1998 MDICP is caused by the relatively low mean parity in the oldest age group. The other age specific mean parities are similar throughout all the survey rounds.

Our comparison of the mean parities for the women who were interviewed in the two respective surveys shows that the overall mean parity (unweighted) of the 1998 MDICP is not substantially larger than that of the 2000 MDHS and that the youngest women in the 1998 MDICP have higher mean parities. The two data sets have similar age distributions and the proportions of women who have ever used contraceptives are also similar. However, the age specific proportions of women with some education in the MDICP are generally higher and we expected this to favour lower mean parities for the women who were interviewed in the MDICP survey.

The age specific mean parities for the 2000 MDHS are generally lower than the parities for the 2001 MDICP, as is the overall mean parity. These observations should be interpreted tentatively because the two data sets are confounded by age. Further, the age specific proportions of women who have ever used contraceptives in the 2001 MDICP are relatively higher. It is likely that the 2001 MDICP mean parities are higher than those of the 2000 MDHS because the MDICP has larger proportions of women in the older reproductive ages.

The mean parities for the women who were interviewed in the 2004 MDICP are generally larger than the parities for the women who were interviewed in the 2004 MDHS. A likely explanation is that the MDICP interviewed larger proportions of older women as shown in Figure 3.4 because the MDICP data set has larger proportions of women who ever used contraceptives.

The mean parities for the women who were interviewed the 2006 and 2008 MDICP surveys are also larger than the parities for women in the 2004 MDHS. We also suspect that the differentials are caused by the relatively larger proportions of older women in the MDICP surveys. The different age distributions are the most likely cause because the proportions of women who have ever used contraceptives are higher for the MDICP data sets relative to the MDHS. Further, the overall proportion of women with some education in the 2006 and 2008 MDICP are larger relative to the proportions for the 2004 MDHS. Hence, the other conditions favour higher parities for the respondents who were interviewed in the MDHS.

In Table 3.16, the weighted mean parities for the latter surveys are substantially smaller than the unweighted parities for the same surveys. Therefore, it is likely that the unweighted MDICP mean parities are larger because they are calculated from data which is confounded by age. We also observe that the weighted parities for all the data sets confirm that the MDICP mean parities are generally larger than the parities for the MDHS as earlier noted. We also observe that the overall parities decline for subsequent MDICP surveys and that the highest parity is for the 2001 data set.

Our investigations have shown that the age distributions of the women who were interviewed in the 2000 MDHS and the 1998 MDICP survey are comparable, while the other paired distributions are different. Despite the larger proportions of women with some education in the 1998 MDICP relative to the 2000 MDHS, and the similar proportions of women who have ever used contraceptives in the two surveys, we still observe larger mean parities for the 1998 MDICP relative to the 2000 MDHS. Hence, it is not clear why the parity differentials are observed for these data sets. The results of our investigations suggest that the longitudinal nature of the MDICP surveys contributes to the larger proportions of older women in the latter MDICP surveys, while the age distributions of the women who were interviewed in the MDHSs have the same general structure. We also noted that, the addition of never-married women to the MDICP age structure in 2004 produced larger proportions of women aged 15-19 years relative to the same proportions for the other MDICP distributions. Further, despite the addition of young women in 2004, the MDICP sample continued to have large proportions of older women in the latter rounds because of the gradual aging of the women initially selected in 1998. Therefore, these older women had a major effect on the fertility differentials that we observed between the women who were interviewed in the MDHSs and the MDICP surveys.

We have found that much of the difference between the MDHS and MDICP data can be attributed to differences in the underlying age distribution; and that Anglewicz, Adams, Obare *et al* had erred in not taking this into account. Thus, their mean parity comparisons are misleading because the two data sets are confounded by age. After controlling for confounding by calculating weighted mean parities, the parities of the MDICP data become marginally larger than those of the MDHS. We suspect that the MDICP mean parities may be inflated because there are very few women in the youngest age groups of the 1998 and 2001 data sets, and the oldest age groups of the 1998, 2001, and 2004 data sets. Therefore, we are going to use the total fertility rate as an overall fertility measure for each data set because it is an age standardised measure which removes the effect of age confounding and the effect of having a few women in some age groups.

4 FERTILITY ESTIMATION

In this chapter, we estimate the fertility implied by the MDICP data. We shall use the consistent parity data and the imputed age data from chapter three. First, we describe how we are going to use the inter-survey parity increment method to estimate fertility. Then, we compare our fertility estimates with the rates implied by the Malawi Demographic and Health Survey (MDHS) data. In the last section, we briefly discuss our findings.

4.1 Fertility estimation method

We are going to estimate fertility by using a variation of the inter-survey parity increment method which is described in United Nations (1983). The original method is mainly used to estimate fertility from surveys which were held five years apart. The method requires the data of the number of women aged 15-49 years (classified by five-year age groups) and the number of children ever-born (classified by the five-year age groups of the mothers). The fertility of each age group of women in the second survey is compared to the fertility of women who are five years younger in the first survey. Thus, the age specific parity increments are measured as the increase in the age specific mean parities between the two time points. Hypothetical cohort parities are calculated by adding the parity increments of all the younger age groups up to the exact age group of interest. The cumulative fertility at each exact age is estimated by applying the polynomial integrals on the hypothetical cohort parities. The integrals of the polynomials have coefficients whose shape defines the cumulative fertility schedule. To obtain annualised age specific fertility rates, the differences of the cumulative fertility of consecutive age groups are divided by five.

Our method uses the total births (classified by the ages of the mothers) and the number of life years lived by the women at each age during an inter-survey period. To calculate the number of inter-survey births by each woman, we subtract her recorded parity in the first survey from her parity in the second survey. This improves the accuracy of our estimates by insuring that the estimated fertility rates are for women who were interviewed in both surveys. The numerators required for deriving the fertility rates are found by counting the number of births by the women at each age during the period and the age specific years of exposure are found by summing the number of years lived by the women at each age during the period. Hence, the age specific fertility rates (ASFRs) are calculated by dividing the age specific births by the age specific life years of exposure. The original method measures the fertility of the synthetic cohorts of women in two consecutive surveys, while our method measures the fertility of real cohorts. A more detailed description of the manipulations required to derive numerators and denominators for our method is given below.

4.1.1 Numerator: classification of births by age of mother at birth

In order to classify the inter-survey births by the age of the mother, we use the rule of thumb that a woman can give birth to a maximum of two children during a three-year period. This is reasonable assumption because the period of breast feeding is generally associated with amenorrhea and according to a study by Zulu (2001), 90 per cent of women in Malawi will still be breast feeding six months after giving birth. Generally, there is a chance of between 95%-98% of a women falling pregnant during the time of breast feeding. Further even though a woman may start ovulating after stopping breastfeeding, during that time, the probability of getting pregnant is low. Hence during a period of 18 months a woman is expected to give birth to one child. In a population in which few women use contraceptives and women breastfeed for a prolonged time, a period of at least six month of infertility is reasonable. Further, the likelihood of having multiple births is a rare event. Therefore, during the 1998-2001 and the 2001-2004 inter-survey periods, each woman is assumed to have given birth to a maximum of two children. During the 2004-2006 and the 2006-2008 periods, each woman is assumed to have given birth to at most one child. Since we do not have some of the interview dates, we shall assume that the interviews were done at the same time of the year during all the survey rounds. The two early surveys only recorded the youngest child's date of birth, while the household rosters of the latter surveys only recorded the data of individuals who usually live in each household. Hence, in order to classify the mothers' ages at birth, we shall assume that births occurred uniformly between the surveys because the birth history data are incomplete.

To classify births by the mothers' ages, we first find the number of inter-survey births by each woman. Then, we determine the age of each woman at the time of the first survey by adding half a year to the recorded age in the first survey. This is done because the mean age of women with the same recorded age (in whole years) is approximately six months (half a year) older than their recorded age. The six months difference is explained by the fact that approximately half of the women had celebrated their birthdays on dates which are within six months before the survey, while the other half had their birthdays between six to twelve months before the survey. Therefore, during the two-year inter-survey periods, each woman will be one year older at the time of her child's birth compared to her adjusted age at the time of the first survey. During the three-year inter-survey periods, each woman will be one year older at the time of her first birth, and two years older at the time of her second birth compared to her adjusted age at the time of the first survey.

In our final step, we derive the number of births by women of the same age during the inter-survey period (B_x). These are found by counting the number of births by women of the same age.

4.1.2 Denominator: Estimation of exposure to the risk

The age specific exposure years are the total number of person years lived by women of the same age during an inter-survey period. As explained in section 4.1.1, women aged x at the time of a survey will on average be aged $x + 1/2$, assuming birthdays are distributed uniformly over the year. This implies that each woman's real age is approximately half a year older than her recorded age and so we estimate the years of exposure which she lives at each age accordingly. Women are exposed, on average from $x + \frac{1}{2}$ to $x + 2\frac{1}{2}$ years for the two-year inter-survey ranges; and to $x + 3\frac{1}{2}$ years for the three-year periods.

To find the age specific exposure years, we sum the years of exposure lived by the women at each age. Hence, the total years of exposure lived by women of the same age during an inter-survey period are calculated by using Equation 4.1.

$$L_x = \sum_{i=1}^{n} P_{x_i}$$ **Equation 4.1**

Where P_{x_i} are the number of exposure years lived by each woman (i) at age x, and L_x are the total years of exposure lived by the women at age x. Therefore, to obtain the ASFRs we divide the age specific births by the total number of life years lived by women of the same age as shown in Equation 4.2.

$$ASFR_x = \frac{B_x}{L_x}$$ **Equation 4.2**

4.2 MDICP fertility rates

Since the 1998 and 2001 MDICP surveys interviewed ever-married women only, the fertility data for the 1998-2001 and the 2001-2004 inter-survey periods are for women who have ever been married. We also know that never-married women were interviewed in 2004 for the first time. Therefore, the data for the 2004-2006 and 2006-2008 periods pertain to the combined samples of ever-married and never-married women. However, Table 3.2 shows that the proportion of never-married women declines from ten per cent in 2004 to only three per cent by 2008. This implies that the effect of pre-marital fertility on the overall fertility of the MDICP sample is expected to decline over time. Based on these observations, we cannot conduct fair comparisons of the fertility of all the women who were interviewed in the five survey rounds because of the variation of the proportion of never-married women who were interviewed in the five rounds. Instead, we shall investigate the fertility of ever-married women only.

4.2.1 MDICP marital fertility rates

In this section, we examine the marital fertility rates of the four MDICP inter-survey periods. We observe the highest TFR for the 1998-2001 period, while the lowest rate is for the 2004-2006 period. The TFR for the 1998-2001 period is substantially higher than the TFR for ever-married rural women who were interviewed in the 2000 MDHS. The observed age specific fertility rates for the youngest women in the early inter-survey periods are lower compared to the corresponding rates for women in the latter periods, while the rates for older women in the latter periods are lower compared to the corresponding rates for women in the early periods. Table 4.1 summarises the observed age specific fertility rates for ever-married women who were interviewed in the 2000 MDHS, 2004 MDHS, and the four MDICP inter-survey periods.

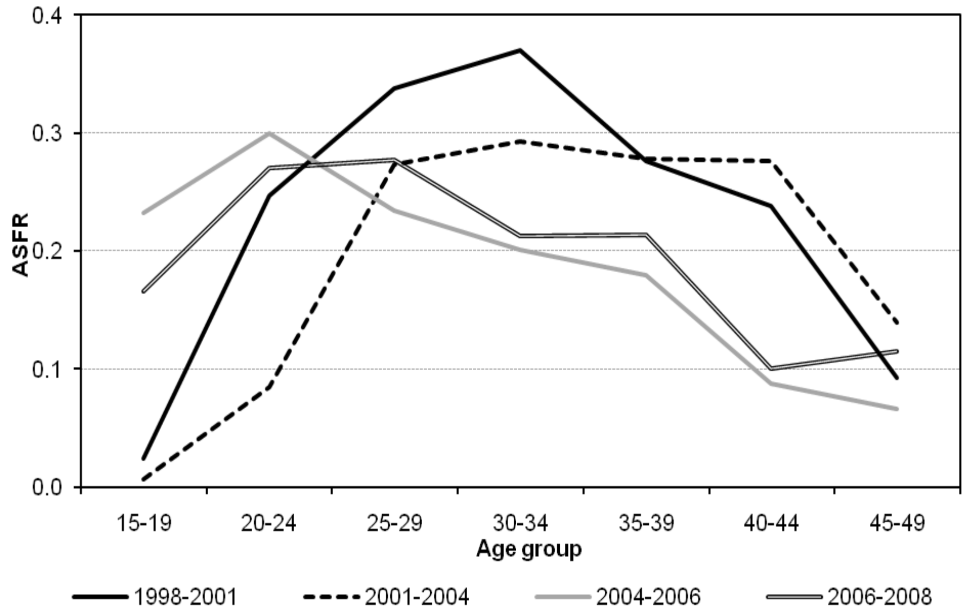
Table 4.1 Observed MDICP marital fertility rates by five year age group and inter-survey period

	MDHS		*MDICP inter-survey period*			
Age group	*2000*	*2004*	*1998-2001*	*2001-2004*	*2004-2006*	*2006-2008*
15-19	0.181	0.222	0.025	0.007	0.233	0.166
20-24	0.162	0.210	0.247	0.085	0.300	0.270
25-29	0.144	0.197	0.338	0.273	0.235	0.277
30-34	0.125	0.184	0.370	0.293	0.201	0.213
35-39	0.107	0.171	0.277	0.278	0.179	0.214
40-44	0.090	0.157	0.238	0.277	0.088	0.100
45-49	0.073	0.143	0.093	0.139	0.066	0.115
TFR	7.2	6.9	7.9	6.8	6.5	6.8

Source: 2000 MDHS, 2004 MDHS 1998 MDICP, 2001 MDICP, 2004 MDICP, 2006 MDICP, and 2008 MDICP

Our investigations show that the fertility rates for the youngest women during the 1998-2001 and the 2001-2004 periods are substantially lower than the corresponding rates for women during the latter periods. In addition, we also note that the age specific rates for middle aged women during the two latter periods are fairly uniform compared to the corresponding rates for the early periods. These observations give the false impression that the TFR was declining, while the fertility of young women was increasing during the ten-year period. We suspect that these changes are caused by data defects because fertility rates do not change considerably over a short period of time. The age distributions of the observed fertility rates for the four inter-survey periods are shown more clearly in Figure 4.1.

Figure 4.1 Observed MDICP inter-survey marital fertility rates, 1998 to 2001, 2001 to 2004, 2004 to 2006, and 2006 to 2008



Source: 1998 MDICP, 2001 MDICP, 2004 MDICP, 2006 MDICP, and 2008 MDICP

The fertility distributions for the 1998-2001 and the 2001-2004 inter-survey periods are suspicious because they are different from the corresponding distributions for the latter periods and the distributions implied by the two MDHS data sets. Contrary to our observations, we expected the highest fertility to be observed at an older age for the latter periods because of the favourable family planning policies which were introduced in Malawi during the latter years. In addition, we are concerned that there are very few women in the youngest age group of the two early inter-survey periods as this could lead to bias. Therefore, our preliminary conclusion is that our fertility estimates for the two early inter-survey periods are unreliable because of data defects, while the estimates for the latter periods are more reliable because they are consistent with the rates implied by the MDHS data. However, we are concerned that the latter inter-survey periods are too short and so they cannot provide enough data to be used to accurately estimate the fertility of each period. Hence, we proceed to investigate the fertility of the two latter periods separately. We begin by describing the Brass polynomial which is used for smoothing an age specific fertility distribution.

4.2.2 Brass polynomial

After estimating the age specific fertility rates by using Equation 4.2, we obtain an approximately concave pattern of the age specific fertility distribution. However, since the MDICP data sets may have very few women in the oldest and youngest age groups, the concave shape may be distorted by random errors. Hence, the Brass polynomial is used to estimate a smoothed fertility distribution from an estimated age specific fertility distribution. The polynomial was developed by Brass (1975) and Hoem, Madsen, Nielsen *et al* (1981). Gage (2001) defines the polynomial as stated in Equation 4.3.

$$m(x) = C(x-d)(d+w-x)^2 \qquad \textbf{Equation 4.3}$$

Where $m(x)$ is the age specific birth rate at age x, C is a measure of the level of fecundity, d is the lower age at fecundity, and w is the length of the reproductive period. We shall estimate C by using the fact that the integral of the age specific fertility distribution defined by $m(x)$ is equal to the observed total fertility rate (TFR) from the data.

Since, $TFR = \int_d^{d+w} m(x)$, we have

$$TFR = \int_d^{d+w} C(x-d)(d+w-x)^2\,dx$$

$$\Rightarrow TFR = C\left(\frac{w^4}{3} - \frac{w^4}{4}\right)$$

$$\Rightarrow C = \frac{TFR}{\left(\frac{w^4}{12}\right)}$$

$$\therefore C = \frac{12 \times TFR}{w^4}$$ **Equation 4.4**

The Brass polynomial is a third degree polynomial. Therefore, the value of d is the solution of the polynomial and w is the minimum turning point. To fit the Brass polynomial, we first generate the fitted age specific fertility rates ($ASFRs_x^{Brass}$). Then, we use the SOLVER function in Microsoft EXCEL to find the values of d and w which minimise the squared differences between the $ASFRs_x^{Brass}$ and the observed ASFRs. These values are used to generate the final estimates of the $ASFRs_x^{Brass}$. To estimate the age specific fertility rates in five year age groups, we sum the $ASFRs_x^{Brass}$ into five year age groups and then we divide each of these sums by five as shown in Equation 4.5.

$${}_5ASFR_x^{Brass} = \frac{\sum_{x=a,5}^{a+5} ASFR_x^{Brass}}{5}$$ **Equation 4.5**

We define a to be a multiple of five from 15 to 45 years. We them calculate the total fertility rate (TFR_{Brass}) from the smoothed rates by multiplying the sum of the (${}_5ASFRs_x^{Brass}$) by five. Equation 4.6 shows the formula which is used to calculate the TFR_{Brass} from the smoothed rates.

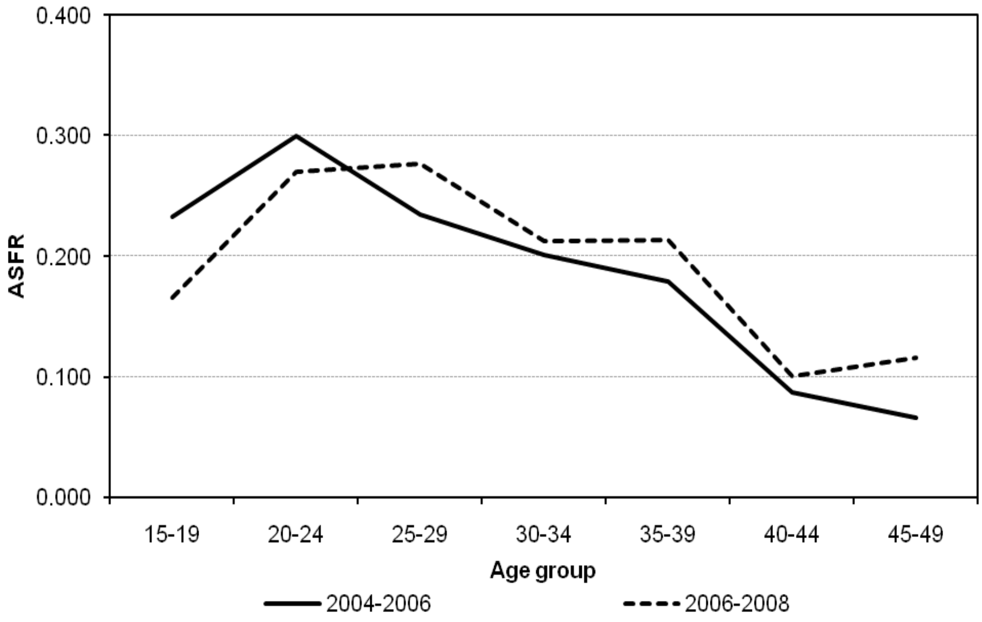
$$TFR_{Brass} = 5 \times \sum_{x=15,5}^{45} {}_5ASFR_x^{Brass}$$ **Equation 4.6**

We shall use the Brass polynomial to smooth the fertility distribution of the 2004-2008 MDICP inter-survey period.

4.2.3 MDICP fertility rates from 2004-2008 and comparison with the 2000 MDHS rates

As mentioned earlier, the fertility distributions of the latter inter-survey periods are similar. The slight differences are probably caused by random data errors.

Figure 4.2 Observed marital fertility rates of the MDICP inter-survey periods, 2004-2006 and 2006-2008

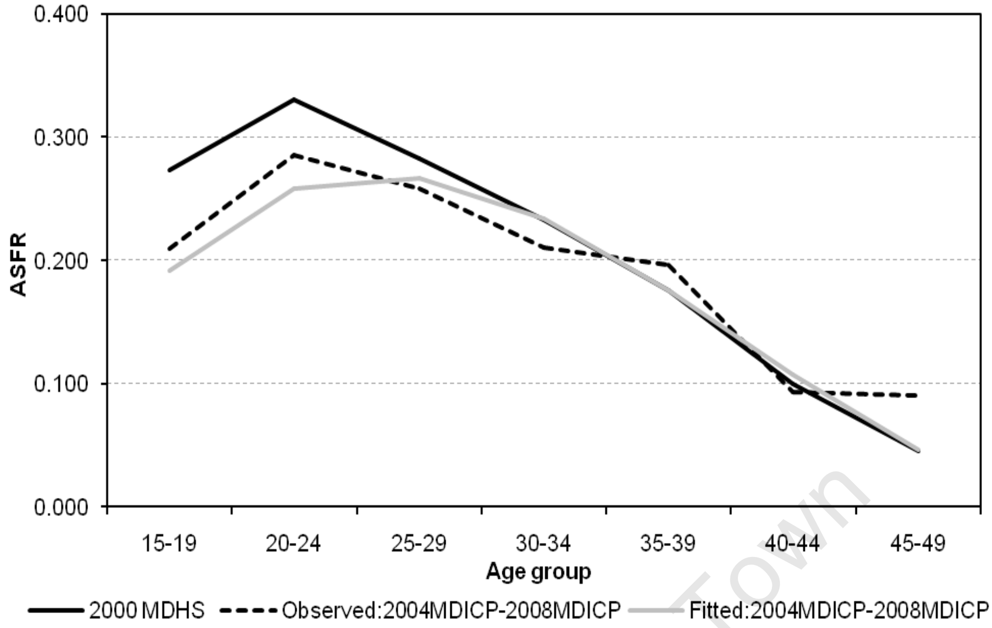


Source: 2004 MDICP, 2006 MDICP, and 2008 MDICP

Further, the differentials arise because each of the inter-survey periods is only two years long. Therefore, the estimated rates will differ because of slight changes to the data. Figure 4.2 shows the age specific fertility rates of the 2004-2006 and the 2006-2008 inter-survey periods. We shall combine the data of the two periods and then compare the fertility of the combined data to the fertility implied by the MDHS data. We estimate the age specific fertility rates for the combined period by adding the age specific births of the two periods. Then, we divide the combined births by the sum of the corresponding age specific exposure years. We smooth the age specific fertility distribution by using the Brass polynomial.

The observed age specific fertility rates for the 2004-2008 MDICP inter-survey period are higher at the oldest age group and they are lower at the youngest age groups compared to the 2000 MDHS rates, while the smoothed rates are similar to the MDHS rates at the oldest age groups. The TFR implied by the MDICP data is slightly lower than that of the MDHS. We also observe that women aged 20-24 record the highest fertility in both data sets. Figure 4.3 and Table 4.2 summarise the age specific fertility rates for the 2004-2008 MDICP inter-survey period and the 2000 MDHS.

Figure 4.3 Comparison of the Observed and Brass fitted marital fertility rates for the 2004-2008 MDICP inter-survey period to the rates for women who were interviewed in the 2000 MDHS



Source: 2004 MDICP, 2006 MDICP, 2008 MDICP, and 2000 MDHS

Table 4.2 Comparison of the Observed and Brass fitted marital fertility rates for the 2004-2008 MDICP inter-survey period to rates for the women who were interviewed in the 2000 MDHS

	MDHS	*2004 MDICP-2008 MDICP*	
Age group	*2000 MDHS*	*Observed:2004-2008*	*Fitted:2004-2008*
15-19	0.273	0.209	0.191
20-24	0.330	0.286	0.258
25-29	0.283	0.259	0.267
30-34	0.233	0.210	0.234
35-39	0.176	0.197	0.176
40-44	0.100	0.093	0.108
45-49	0.045	0.091	0.046
TFR	7.2	6.7	6.4

Source: 2000 MDHS, 2004 MDICP, 2006 MDICP, and 2008 MDICP

Our investigation confirms that the TFR of the MDICP data from 2004-2008 is broadly consistent with the rate of the 2000 MDHS because of the similar age specific fertility distributions. However, the MDICP distribution is substantially lower at the youngest ages. Further, the combined trend of the fertility rates of the two respective surveys shows that the fertility of ever-married rural women declines during the period from 2000 to 2008. We also observe that both the MDHS and the latter MDICP surveys have higher fertility for women who are aged 15-24 years and lower fertility for older women compared to the corresponding rates of the early MDICP surveys.

4.3 Summary of results and the possible explanations

In Table 4.1, we observed that the women aged 20-24, who were interviewed during the latter MDICP surveys and the two MDHSs record the highest age specific fertility rates, while women aged 30-34 record the highest rates during the early MDICP surveys. This gives the impression that the age specific fertility rates for young women in the MDICP data sets were increasing, while the rates for older women were decreasing during the ten-year period. The age distribution of fertility is not prone to rapid changes over a short period of time. Thus, the distinctively shaped marital fertility distributions of the early inter-survey periods are probably caused by errors in the data. Our conclusion is based on the fact that the fertility distributions of the early periods are not consistent with the distributions for the latter periods and those of the 2000 MDHS data. Therefore, we are satisfied that fertility rates are broadly consistent for the latter periods and we cannot draw meaningful conclusions from the rates of the early periods.

Our fertility estimates in Table 4.1 show that the total fertility rates for ever-married women in the 2000 and 2004 MDHS data sets are slightly higher than the corresponding rates for the 2004-2006 and the 2006-2008 inter-survey periods. In both surveys, women aged 20-24 have the highest fertility. In addition, the fertility distributions of the two respective surveys confirm that the fertility of ever-married rural women was higher during the time of the 2000 MDHS and that it declined marginally during the latter years. This is shown by the lower fertility estimates for the 2004-2008 inter-survey period. Figure 4.3 shows that the age specific fertility rates of the 2004-2008 MDICP inter-survey period are similar to those of the 2000 MDHS and that the fertility rates for young women are lower for the MDICP data set. Therefore, the lower overall fertility implied by the MDICP data is most likely to be caused by the lower fertility of the young women who were interviewed in the MDICP surveys.

5 CONCLUSIONS

The fertility transition in Malawi was probably delayed by the fact that the government did not allow women to use modern family planning methods. Further, the results of earlier studies on MDHS data suggest that the rate of fertility decline for the women who live in rural areas is much slower than that of urban women. Although the MDICP data sets have missing fertility data, we used the inter-survey parity increment method to estimate the fertility of the women who were interviewed in the surveys. We have confirmed that the MDICP and MDHS fertility data are confounded by age, and that the data of the early MDICP surveys produce misleading fertility indicators. In this chapter, we discuss these findings.

5.1 MDHS and MDICP sample differentials, and MDICP data quality

The analysis by Anglewicz, Adams, Obare *et al* suggests that the parity differentials between the women who were interviewed in the MDHSs and the MDICP surveys are caused by the fact that the two surveys interviewed women who were sampled from different subsets of the rural population. Although our investigations could not establish the effect of the different samples, we found that there is no substantial difference between the fertility rates for the women who were interviewed in the latter surveys. This may indicate that the effect of sampling different subsets of the rural population is not a major causal factor of the observed fertility differentials. However, we have noted that the uniqueness of the MDICP districts is caused by the fact that the selection of the MDICP sites had a particular purpose and therefore the respondents who were interviewed in each site had unique characteristics. For example, we know that Balaka district was chosen in order to follow the results of a community based initiative by GTZ and it was chosen instead of Chiradzulu district (southern region) because it contained a large proportion of the Yao people who are predominantly Muslim. In addition, the MDICP sites are predominantly rural while the rural MDHS sample consists of townships and semi-urban areas. Therefore, these factors are likely to contribute to the higher fertility which is observed for the MDICP data.

The mean parities for the two data sets were expected to differ because the MDICP is a longitudinal study, while the MDHS is a cross sectional study. Therefore, the age distributions of the women who were interviewed in the latter MDICP surveys were expected to have larger proportions of older women because of the aging of the women who were initially interviewed in the early survey rounds, while the age distribution of women who were interviewed in each of the MDHSs maintains a more standard structure closely resembling the actual population structure of Malawi. Our data analysis showed that the age distributions of the women who were interviewed in the latter MDICP surveys are indeed affected by the aging of the initial respondents who were interviewed in 1998 and the addition of adolescent women in 2004. Therefore, aging contributed to the confounded data.

Although the early MDICP surveys interviewed ever-married women only, while never-married women were included in the study for the first time in 2004, the proportion of never-married women declines substantially in each of the subsequent surveys after the 2004 survey round. Therefore, it is difficult to conduct a fair comparison of the fertility of all the women (ever-married and never-married) who were interviewed in consecutive MDICP survey rounds. Further, it is also difficult to estimate the fertility rates implied by the MDICP data sets because there are a lot of women with unrecorded ages.

The early MDICP surveys only recorded the year of birth, while the latter surveys recorded the age of each woman. In addition, the MDICP surveys did not record the dates of birth of each woman's children or indeed the mother's date of birth either. Hence, the unavailable dates could have compromised the accuracy of our fertility estimates because we had to resort to using the inter-survey parity increment method to estimate fertility. Generally, our investigations of the determinants of fertility were hindered by the erratic recording of the relevant data.

The 2008 MDICP surveys did not collect data on contraceptive use, while the 2004 data set combines the respondents with no education and those who have only primary education into one category. This complicated our investigation of these two variables over time. Further, unlike the MDHS data sets, some variable names in the MDICP data sets differ in each survey round. Hence, a lot of time was spent identifying each variable in each of the data sets.

5.2 Findings

We have confirmed that the fertility data of the women who were interviewed in the MDHSs and the MDICP surveys are confounded by age. Hence, the comparisons by Anglewicz, Adams, Obare *et al* are biased. We also noted that the latter MDICP data sets have larger proportions of older women because of the gradual aging of the women who were initially interviewed in the early survey rounds. Both these reasons explain why higher overall mean parities are observed for the women who were interviewed in the MDICP surveys relative to the parities for the women who were interviewed in the MDHSs. These reasons also explain why higher mean parities are observed for the women who were interviewed in the latter MDICP surveys relative to the parities for the early rounds.

We controlled for age confounding by calculating weighted mean parities which we weighted by the age distributions of the women who were interviewed in the MDHSs. We observed that the mean parities for the women who were interviewed in the MDICP surveys became marginally larger than the mean parities for the women who were interviewed in the MDHSs. These differentials are caused by the small number of women in the youngest and oldest age groups of the MDICP data sets. In order to determine if our fertility estimates conformed to the fertility implied by MDHS data, we went on to estimate the fertility rates implied by the MDICP data.

We discovered that the age distributions of the fertility rates of the two early MDICP inter-survey periods are uncharacteristically high for middle-aged women, while the rates for the latter periods and those for the MDHS data sets are high for younger women. Hence, the fertility rates of the two early inter-survey periods are inconsistent with the rates of the latter surveys, and the rates of the more reliable MDHS data sets. Thus, we concluded that the fertility data of the early MDICP surveys produce misleading results.

Despite the concerns noted above, our comparison of the fertility rates of the latter MDICP inter-survey periods to rates implied by the MDHS show that the fertility rates of the ever-married women in the MDICP surveys are slightly lower than the rates of ever-married rural women who were interviewed in the MDHSs of 2000 and 2004. Our observations are consistent with the notion by Mijoni (2005) that the rate of fertility decline for rural women in Malawi is slow.

In summary, we cannot fully study the fertility trend of the women who were interviewed in the MDICP surveys because the fertility data of the early surveys are biased. Hence, it is difficult to conduct a fair comparison of the fertility of all the women who were interviewed in the MDICP surveys to the fertility of the women who were interviewed in the MDHSs. Further, since we have found that the MDHS and MDICP data are confounded by age, this means that any comparison of the two data sets must control for this confounding. We also recommend that any user of the

MDICP data should check the data for inter-survey consistency before conducting their research.

5.3 Limitations of our investigation

The source of our problems with the MDICP data is that, unlike the MDHS, the MDICP was not primarily designed to collect fertility data. Therefore, it may not provide the appropriate data for us to thoroughly interrogate the fertility of the women who were interviewed in the two surveys.

There are only two available MDHS data sets, while there are five MDICP data sets. In addition, the 2008 MDHS data set has not yet been released. Therefore, we could only compare the data of the 1998 and 2001 MDICP surveys to the data of the 2000 MDHS, and the 2004, 2006 and 2008 MDICP data sets to that of the 2004 MDHS. Further, the 1998 MDICP survey was held two years before the 2000 MDHS, while the 2006 and 2008 MDICP surveys were each held at least two years after the 2004 MDHS. Therefore, our findings from our comparisons of the MDHS and MDICP data sets may be compromised by the fact that the compared surveys were not held during the same year.

We cannot correct the parity data of each woman to become consistent throughout all the five survey rounds because we cannot determine the survey in which it was incorrectly recorded. In addition, there are a lot of women who have missing parity data in the respective surveys. Thus, the accuracy of our fertility estimates could be compromised. The only way to offset this error is to improve the effectiveness of the data collection and recording.

Since the MDICP is a longitudinal study, the total number of women who were interviewed in each of the subsequent surveys declined due to attrition and a lot of women were not interviewed in consecutive survey rounds. Therefore, the inter-survey parity increment method may not accurately estimate the fertility of the women who were interviewed in the MDICP surveys. Further, the method does not consider the fertility of women who move into the districts during an inter-survey period.

Just like the standard procedure which is used to estimate the fertility of MDHS data, the inter-survey parity increment method may produce biased results because it assumes that the fertility of the women who migrated or died is the same as the fertility of the women who were interviewed. This assumption may not hold for the MDICP data because some other studies show that the women who are lost to attrition in the MDICP study have unique characteristics. For example, Reniers (2003) had found that the women who were lost to attrition between the MDICP surveys of 1998 and 2001 had fewer children, and were less likely to be members of indigenous churches compared to those who were interviewed in both surveys

According to United Nations (1983), there is a tendency by older women to omit children who died or migrated when reporting their parity. Therefore, the fertility of older women must be interpreted with caution especially if it is substantially lower than that of younger women. The MDHS data are subject to the same error, albeit to a lesser extent since the birth history data are meticulously verified.

When we used the inter-survey parity increment method, we used the rule of thumb that each of the women could give birth to a maximum of two children during a three-year period. When we used this assumption we disregarded the fact that there were some women who gave birth to twins or triplets. Therefore, our fertility estimates may be lower than the estimates that would arise if we consider multiple births. However, the effect of this error is negligible because multiple births are rare.

5.4 Conclusions and Recommendations

We have found that the MDICP and MDHS parity data are confounded by age. We went on to estimate the fertility rates implied by the MDICP data and to compare the estimated rates with the rates implied by the MDHS data. We found that the fertility rates of the early inter-survey periods are inconsistent with the rates of the other surveys and so we concluded that the fertility rates of the early surveys are biased. We could not estimate fertility by using the direct method of dividing the births by the exposure years lived because the MDICP data sets do not provide the relevant data.

Despite the inconsistent data, we observed that the overall fertility rates implied by the 2004, 2006, and 2008 MDICP data sets are marginally lower than those of the 2000 and 2004 MDHS data sets. The difference is caused by the low fertility estimates for the women who were younger than 25 years in the MDICP surveys relative to the corresponding fertility for the women who were interviewed in the MDHSs. This indicates that the marginal difference between the overall fertility rates of the two respective surveys is caused by the lower fertility of the young women who were interviewed in the MDICP surveys. Hence, fertility decline in Malawi can be explained by the decline of the fertility of young women for the reason that the fertility of older women does not show a substantial change during the study period. Due to the fact that never-married women were only included in the MDICP surveys for the first time in 2004, we could not investigate the fertility of all the women who were interviewed in both surveys. We hope to be able to carry out such a research provided the data becomes available in future survey rounds.

Another possible area of future research is to interrogate the MDICP fertility according to the women's background characteristics—for example, the regional comparison of fertility and contraceptive use. This suggestion is based on the fact that the three MDICP districts have unique traditional practices and Doctor(2005) noted that the level of education differs among these groups.

We encountered problems with the MDICP data sets because the dates of important events are not recorded precisely, and some variable names differ in the five data sets. Therefore, before interrogating the fertility implied by the MDICP data, we had to derive the missing information from the data recorded in the other survey rounds. Our fertility estimates could have been more accurate if the date of birth for each woman and her children were recorded in a format that included the day, mouth and year of birth. Further, the MDICP data sets must maintain the same format of naming variables so that it easy to identify each variable in all the data sets. However, we must note that the quality of the MDICP data was probably also affected by the fact that there was no time to edit the data as is the case with the DHS data because the MDICP survey were held two or three years apart. Further, the data quality was also affected by illiteracy—particularly amongst rural women as is the case in most studies. Therefore, the uncorrected errors have a consequence on any comparisons that can be made with other data set.

The MDICP is a new study and it provides opportunities of research in the other areas of demography. However, the quality of the data needs to be improved so that it can cater for such research. Our investigations highlight the potential sources of bias which may affect the results of any research using MDICP data. Thus, we suggest that, the accuracy of time related information can be improved by using the history of calendar events or the respondents may be asked to state the number of years prior to the interview that an event occurred. We also suggest that when respondents know the exact date of events, the full date should be recorded although this is not feasible when conducting interviews in developing countries because of the low levels of literacy.

REFERENCES

Anglewicz, P., Adams, J., Obare, F., Kohler. H.P. and Watkins, S. 2009. "The Malawi Diffusion and Ideational Change Project 2004-2006: Data collection, data quality, and analysis of attrition", *Demographic Research* **20**(21):503-540.

Bangala La Mtsogolo 2008. *BLM 2008 Annual Report: Providing Choices in Reproductive Healthcare Countrywide*. Blantyre: Banja La Mtsogolo.

Bignami-Van Assche, S., Reniers, G. and Weineb, A.A. 2003. "An Assessment of the KDICP and MDICP Data Quality: Interviewer Effects, Question Reliability and Sample Attrition", *Demographic Research* **Special Collection 1**(Article 2):31-76.

Bongaarts, J. and Potter, R.E. 1983. *Fertility, Biology, and Behaviour: An Analysis of Proximate Determinants*. New York: Academic Press.

Bongaarts, J., Odile, F. and Lesthaeghe, R. 1984. "The Proximate Determinants of Fertility in sub-Saharan Africa", *Population and development Review* **10**(3):511-537.

Brass, W. 1975. *Methods for Estimating Fertility and Mortality from Limited and Defective Data,* Chapel Hill: International Program of Laboratories for Population Studies, Department of Biostatistics, School of Public Health, Carolina Population Centre, University of North Carolina.

Chimbwete, C., Watkins, S.C. and Zulu, E.M. 2005. "The evolution of population policies in Kenya and Malawi ", *Population Research and Policy Review* **24**(1):85-106.

Doctor., H.V. 2005. "Women's schooling and religious affiliation in Malawi at the end of the twentieth century", *International Journal of Educational Development* **25**(5):481-492.

Gage, T.B. 2001. "Age-Specific Fecundity of the Mammalian Populations: A Test of the Three Mathematical Models", *Zoo Biology* **20**(6):487-499.

Garenne, M.M., Tollman, S. and Khan, K. 2000. "Premarital Fertility in Rural South Africa: A Challenge to Existing Population Policy", *Studies in Family Planning* **31**(1):47-54.

Garenne, M.M. 2008. *Fertility Changes in sub-Saharan Africa*. DHS Comparative Reports No. 18. Calverton, Maryland, USA: Macro International lnc.

Harwood-Lejeune, A. 2000. "Rising Age at Marriage and Fertility in Southern and Eastern Africa", *European Journal of Population* **17**:261-280.

Hoem, J.M., Madsen, D., Nielsen, J.L., Ohlsen, E., Hansen, H.O. and Rennermalm, B. 1981. "Experiments in Modelling Recent Danish Fertility Curves", *Demography* **18**(2):231-244.

Khan, S., Mishra, V., Arnold, F. and Abderrahim, N. 2007. *Contraceptive Trends in Developing Countries*. DHS Comparative Report No. 16. Calverton, Maryland, USA: Macro International lnc.

Kohler, H.P., Taulo, F., Masanjala, W., Watkins, S.C. and Behrman, J.R. 2009. *The Malawi Longitudinal Study of Families and Health*. University of Pennsylvania. http://www.ssc.upenn.edu/~hpkohler/malawi/2009-02-mlsfh-datacollection.pdf. Accessed: 17 April 2010.

Kravdal, O. 2001. *The Importance of Education for Fertility in sub-Saharan Africa is Substantially Underestimated When Community Effects are Ignored*. Oslo: Norway: University of Oslo, Department of Economics.

National Statistical Office [Malawi], and ORC Macro. 2005. *Malawi Demographic and Health Survey 2004*. Calverton, Maryland, USA: Macro International lnc.

Mijoni, B.A. 2005. "Examination of the fertility measures from the 1992 and 2000 Malawi Demographic and Health Surveys, *a review of declining fertility*." Unpublished MPhil in Demography Thesis, Cape Town: University of Cape Town.

Miller, K., Zulu, E.M. and Watkins, S.C. 2001. "Husband-Wife Survey Responses in Malawi", *Studies in Family Planning 2001* **32**(2):161-174.

Reniers, G. 2003. "Divorce and remarriage in rural Malawi", *Demographic Research* **Special collection 1**(Article 6):176-202.

Solo, J., Jacobstein, R. and Malema, D. 2005. *Repositioning family planning-Malawi case study: Choice, not chance.* New York: The ACQUIREProject/EngenderHealth.

United Nations Department of International Economic and Social Affairs. Population Division, National Research Council (U.S). Committee on Population and Demography 1983. *Manual X: Indirect Techniques for Demographic Estimation.* New York: United Nations.

Watkins, S.C., Zulu, E.M., Kohler, H.P. and Behrman. J.R. 2003. "Social Interactions and HIV/AIDS in Rural Africa", *Demographic Research* **1**(1).

Weinreb, A. 2007. *Description of 2004 MDICP adolescent sampling process.* http://www.malawi.pop.upenn.edu/Level%203/Papers/PDF-files/Description2004AdolescentSample.pdf. Accessed: 17 April 2010.

Zulu, E.M. 2001. "Ethnic Variations in Observance and Rationale for Postpartum Sexual Abstinence in Malawi", *Demography* **38**:467-479.

APPENDIX

Table A.1 Observed and Brass fitted age specific fertility rates, 2004-2008 MDICP inter-survey period

Age	*Observed: 2004-2008*	*Brass Fitted: 2004-2008*	*Squared Difference*
15	0.000	0.148	0.022
16	0.167	0.173	0.000
17	0.163	0.194	0.001
18	0.430	0.213	0.047
19	0.286	0.229	0.003
20	0.372	0.242	0.017
21	0.286	0.252	0.001
22	0.249	0.261	0.000
23	0.278	0.266	0.000
24	0.244	0.270	0.001
25	0.218	0.271	0.003
26	0.308	0.271	0.001
27	0.268	0.269	0.000
28	0.225	0.265	0.002
29	0.275	0.260	0.000
30	0.197	0.253	0.003
31	0.211	0.245	0.001
32	0.206	0.235	0.001
33	0.245	0.225	0.000
34	0.191	0.214	0.001
35	0.228	0.202	0.001
36	0.163	0.189	0.001
37	0.255	0.176	0.006
38	0.136	0.163	0.001
39	0.201	0.149	0.003
40	0.066	0.135	0.005
41	0.129	0.121	0.000
42	0.100	0.107	0.000
43	0.121	0.094	0.001
44	0.051	0.081	0.001
45	0.133	0.068	0.004
46	0.055	0.056	0.000
47	0.103	0.045	0.003
48	0.084	0.035	0.002
49	0.078	0.026	0.003
TFR	6.7	6.4	

The Brass fitted rates were estimated by using the Brass polynomial:

$$m(x) = C(x-d)(d+w-x)^2$$

Where, C is a measure of the level of fecundity, d is the lower age at fecundity, and w is the length of the reproductive period. $C = \frac{12 \times TFR}{w^4}$. Our estimate of d is 10.6 years and the length of the reproductive life (w) is 44.0 years.