Working paper

A Bayesian model for the reconstruction of education- and age-specific fertility rates: An application to African and Latin American countries

Afua Durowaa-Boateng [1*](afua.durowaa-boateng@oeaw.ac.at)
Dilek Yildiz [1,2] (yildiz@iiasa.ac.at)
Anne Goujon [1,2] (goujon@iiasa.ac.at)

WP-23-007

[1] Wittgenstein Centre for Demography and Global Human Capital, Vienna Institute of Demography of the Austrian Academy of Sciences (VID/ÖAW, IIASA, University of Vienna)
[2] Population and Just Societies (POPJUS) Program, International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria.
*Corresponding author

Approved by:

Name Albert van Jaarsveld
Program: Director General and Chief Executive Officer
Date: 06 July 2023

Table of contents

Abstract

Consistent and reliable time series of education- and age-specific fertility rates for the past are difficult to obtain in developing countries, although they are needed to evaluate the impact of women's education on fertility along periods and cohorts. In this paper, we propose a Bayesian framework to reconstruct age-specific fertility rates by level of education using prior information from the birth history module of the Demographic and Health Surveys (DHS) and the UN World Population Prospects. In our case study regions, we reconstruct age- and education-specific fertility rates which are consistent with the UN age specific fertility rates by four levels of education for 50 African and Latin American countries from 1970 to 2020 in five-year steps. Our results show that the Bayesian approach allows for estimating reliable education- and age-specific fertility rates using multiple rounds of the DHS surveys. The time series obtained confirm the main findings of the literature on fertility trends, and age and education specific differentials.

About the authors

Afua Durowaa-Boateng is a researcher at the Vienna Institute of Demography of the Austrian Academy of Sciences, and a PhD student in Demography at the University of Vienna. (Contact: afua.durowaa-boateng@oeaw.ac.at)

Dilek Yildiz is a research scholar at Population and Just Societies Program at IIASA. She is also a research scholar at Vienna Institute of Demography of the Austrian Academy of Sciences. (Contact: yildiz@iiasa.ac.at)

Anne Goujon is Program Director of Population and Just Societies Program at IIASA. (Contact: goujon@iiasa.ac.at)

Acknowledgements

Introduction

The existence of a relationship, mostly negative, between women's education and fertility has been shown in many settings (e.g., Ahuja and Filmer 1995, Basu 2002, Bongaarts 2010, Brand and Davis 2011, Lutz and KC 2011, Martin and Juarez 1995, Weinberger et al. 1989). The potential mechanisms at play are many, such as the fact that women with higher levels of education spend more years in school, which reduces their exposure to marriage and pregnancy at a younger age. They also tend to have better knowledge and access to contraceptives to control their own fertility (e.g., Bongaarts 2010; Gebreselassie and Shapiro 2016). For example, in Latin America, Weinberger et al. (1989) confirm that women with higher levels of education had fewer children and that 40-67% (overall in Peru, Colombia, Dominican Republic and Ecuador) of the observed decline in fertility rates between the 1970s and 1980s were due to improvements in education. Bongaarts (2010) also explains the negative effects of education on fertility rates in sub-Saharan Africa by the greater demand and use of modern contraceptives by women with higher levels of education. However, many unknowns remain about the precise pathways by which education affects fertility levels in different contexts and at different times. One of the main challenges in studying the relationship between education and fertility is the lack of comparable, consistent, and unbiased data on education- and age-specific fertility rates (EASFR), particularly in low-income countries.

Due to the limited availability of comprehensive and reliable registration systems, the main sources of data on fertility rates in high-fertility countries are various surveys (e.g, Demographic and Health Survey (DHS), Service Provision Assessment (SPA) survey, HIV/AIDS Indicator Survey (AIS), Malaria Indicator Survey (MIS)) and the indirect estimates from the United Nations World Population Prospects (UN WPP 2022a) that use multiple data sources. The total fertility rates (TFR) and age-specific fertility rates (ASFR) published by the UN WPP on a regular basis are widely used. These estimates are consistent and allow for comparisons of fertility over time and across countries; as such they are considered trustworthy. However, these rates are not disaggregated by level of educational attainment. The Wittgenstein Center for Demography and Global Human Capital (WIC) publishes population projections and estimates by level of education based on data collected from multiple sources, including the UN WPPs (WIC 2018), and assumptions about the future. However, the education- and age- specific fertility rates for past years are not estimated. Demographic surveys provide fertility rates for population sub groups, including levels of educational attainment. Among them, Demographic and Health Surveys (DHS) are the main and only source of fertility data in many low-income countries. However, like any survey, they can be subject to different levels of sampling error that often lead to inconsistencies, limiting comparisons across countries and for the same country over time. Furthermore, they are mostly not conducted at regular intervals, such as every five years. The most common inconsistencies have been attributed to birth omissions and displacements (e.g., Al Zalak and Goujon 2017; Pullum and Becker 2014; Schoumaker 2011, 2014). Schoumaker (2014) in an assessment of DHS data categorized most Latin American countries as having "good" or "moderate" data quality when available while many African countries were categorized as having "poor" data quality, reporting different rates for the same periods across consecutive surveys (e.g., Ethiopia). Figure 1 presents the age-specific fertility rates of women with secondary education in the 1990-1995 period estimated by different rounds of the DHS surveys using the DHS birth history module. In countries with good data quality (e.g, Colombia), age-specific schedules are more complete and estimates are relatively consistent for the same years, even for different surveys (see Figure 1).
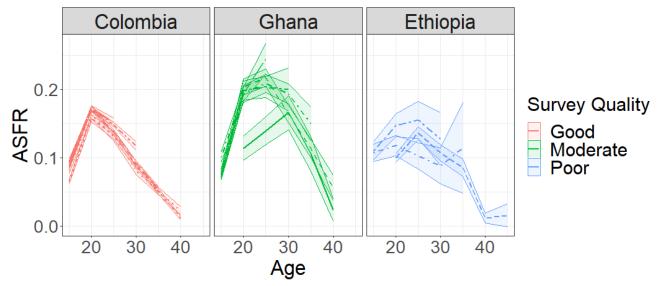
Figure 1: Quality of data for the 1990-1995 period for secondary education ASFR by different DHS surveys

Regarding the lack of time series in demographic surveys, Schoumaker (2013) proposed a Poisson regression model to reconstruct both ASFR and TFR from different rounds of DHS surveys. The method estimates past fertility rates using the birth history module of the survey. While it fills the gaps, the estimates are not consistent with the UN WPP which are used widely (and considered more reliable).

In this research we combine four datasets (UN WPP, WIC, estimates from Yildiz et al. (2023), and DHS) to produce reliable and consistent EASFRs over time and across countries, which are compatible with the UN WPP ASFR, using a Bayesian model.

We follow Bijak and Bryant's (2016) recommendation to employ Bayesian methods in situations where sample sizes are too small, and the data quality is limited or poor. The Bayesian approach has been used in recent years in demographic estimates and population projections (e.g., Alkema et al. 2011; Bryant and Graham 2013; Ellison et al. 2020; Hilton et al. 2019), including by the United Nations (UN) in recent rounds of population projections (Ševčíôková et al. 2018). The modeling carried under this work aims at developing a consistent and reliable dataset that will allow for a better understanding of the impact of education on fertility.

We estimate age specific fertility rates for 41 DHS African and 9 Latin American countries by four levels of education between 1970-1975 and 2015-2020. Our paper contributes to the literature by proposing an advanced statistical model which fills the gap in the time series when data are missing, and by providing complete and UN WPP consistent EASFRs for all 50 countries. We focus on Latin American and African countries in this analysis since they lack, to varying degrees, detailed, regular, and consistent data on EASFR for past years. Moreover, these two regions are interesting as the timing and pace of the demographic transition are different. We plan to extend the research to more regions in the Global South in the future.

Data sources

We use four different data sources for this study: DHS, UN WPP (2022b), UN-consistent education specific TFR from Yildiz et al. (2023), and WIC (2018). A total of 217 DHS conducted in Africa and Latin America are pooled for the analysis (see the full list of countries and surveys in Figure 2 and in Appendix A). Fertility rates are obtained from individual recode datasets in the DHS database (ICF 2004-2017). We consider four levels of education: "No Education", "Primary Education", "Secondary Education" and "Higher Education". These

educational categories are collected in DHS surveys for each country and survey wave and are based on the highest level of education reported by the woman (The DHS program n.d).

Figure 2: Countries and surveys collected, Africa and Latin America

Our analysis uses women's retrospective birth histories as collected in DHS rounds, focusing on reproductive ages between 15 and 49 years. Since our goal is to reconstruct past fertility rates by level of education, we focus on birth histories as far as 30 years before the survey for Latin American countries. For African countries, the birth histories are collected for a shorter period of 15 years, given the lower quality of estimates due to long recalling periods. The retrospective fertility rates by five-year age groups and five-year periods are obtained using the Stata "tfr2" module which provides fertility rates close to those published in the DHS reports (Schoumaker 2013). We utilize all available surveys with the exception of those conducted after 2020.

The second source of fertility data in our research is the 2022 edition of the UN WPP (2022b) which provides ASFRs for five-year periods. The UN WPP collects fertility data preferably on live births by age of mother from civil registration systems as well as all available data, including from the DHS. The UN WPP is updated every couple of years to achieve consistency over time and across different demographic statistics and include new data sources which were previously not included. This process makes the UN WPP a reliable source for global demographic data.

Further, Yildiz et al. (2023) estimated education specific total fertility rates for Sub-Saharan African countries from 1980 to 2015 by 5-year period using multiple data sources including WIC and UN WPP. The authors use a flexible hierarchical Bayesian model that allows education specific estimates to vary with regards to the degree of consistency with the UN data. The estimates are provided for Sub-Saharan African countries only and focus on total fertility rates by education but not by age.

In addition to fertility rates, the size of the female population by level of education for the countries under consideration is obtained from the WIC data explorer (WIC 2018) through the 'epop' function in the "wcde" R package (Abel 2021). These estimates are constructed using an iterative multi-dimensional cohort-component reconstruction model (IMCR) based on historical data on education and mortality (see Speringer et al. (2019) for more details). All data in our analysis concern women aged 15-49 years between 1970 and 2020.

Methodology

Various approaches have been used within the Bayesian framework to estimate fertility rates. For example, Ellison et al. (2020), inspired by the Lexis diagram, developed a Bayesian technique to estimate cohort fertility where births were modeled to follow a Poisson distribution. Earlier, Bryant and Graham (2013) had modeled births as part of a sub-national population estimation model, in which births follow a Poisson distribution centered around expected births at the end of the period, similar to Ellison et al. (2020). As mentioned above, Yildiz et al. (2023) used a flexible hierarchical Bayesian technique to estimate education specific total fertility rates for Sub-Saharan countries for the period 1980 to 2015 using existing sources. Finally, Alkema et al. (2011) estimated total fertility rates for 196 countries by considering their fertility transition phases. Each transition phase was modeled separately taking into account the rate of decline. During the transition period, the TFR was modeled as the previous fertility rate minus the expected five-year decrement, plus an error term. The five-year decrement term follows a defined function while the error term follows a set of normal distributions when a set of conditions are met. For post-transition countries, the fertility rates were modeled as a normal distribution of a first order autoregressive time series model.

Our modeling framework consists of two main steps applied separately to African and Latin American countries. We improve and complement previous Bayesian approaches which did not integrate education and its impact on fertility rates, and did not estimate long time series of age-specific fertility rates. The first step in our approach is to estimate EASFRs for all DHS countries which enter the Bayesian model as initial values. To achieve this, we employ a generalized linear model (glm) with a Poisson link. The data for the glm model are the EASFRs obtained by the STATA tfr2 module using DHS birth histories. We adopt this approach to account for gaps in DHS estimates. We treat the glm estimates as the "input" dataset for our Bayesian model.

The predicted estimates use information from the variables that influence fertility rates in our dataset, including the country itself, the number of women under consideration by five-year periods and age groups, and their education levels. In the event that a whole period schedule is missing (e.g., 2015-2020 period for Latin American countries), the estimates use the aggregate effects of the period in question, the age groups and the country estimates (e.g., 2015-2020 period effect from all countries, educational effect, age group effects, country effects

and the specific interaction effects). To a large extent, the model learns from the other countries in the region since it uses all available information in the dataset to make estimates for the missing values. The glm model involves interaction terms between variables based on the assumption that the effects of these variables are not constant. The regression model for EASFR is defined as:

$$EASFR^{DHS} \sim Education + Age\ Group + Country + Period + Age\ Group * Country +$$
$$Country * Period + Education * Age\ Group + Period * Age\ Group. \qquad (1)$$

Using the glm model, we estimate initial values of the EASFR between 1970-1975 and 2015-2020 in five-year intervals. The estimates produced by the model logically do not match with the UN WPP ASFRs. Also, in the event of missing or poor-quality data, the model estimates higher values than expected by the trend, for example, creating an abrupt halt or reversal of the fertility decline during the missing period. To address these issues and ensure consistency with the UN WPP ASFR, we employ a Bayesian framework in the second step, as explained in the next paragraphs.

The first level (Level 1) starts with calculating the ASFRs as a weighted total of the predicted EASFRs in Equation 2. The weights, w_{cyae}, are the ratio of the population of women in age group a by level of education e to the total population of women aged 15-49 years for each five-year period y in country c according to WIC. Then, in Equation 3, the ASFRs are adjusted to the UN WPP ASFRs for each country for all five-year periods making use of the precision parameter τ^{asfr}, the inverse of the variance. The variance parameter μ^{asfr} is data-driven and reflects the uncertainty in UN WPP ASFRs. In other words, it is used to define the degree of consistency with the UN WPP ASFRs. Since the UN WPP ASFRs are themselves dependent on DHS data, we only include ASFRs from UN WPP to avoid repeating the same information in our Bayesian model. An exercise to investigate the sensitivity of the estimates to this parameter is presented in Appendix B.

Level 1:
$$ASFR^*_{cya} = \sum_{e=1}^{4} \ (EASFR^*_{cyae} \times w_{cyae}) \qquad (2)$$
$$ASFR^{UN}_{cya} \sim N(ASFR^*_{cya}, \tau^{asfr^2}) \qquad (3)$$
$$\tau^{asfr} = 1/\sigma^{asfr} \qquad (4)$$
$$\sigma^{asfr} = \mu^{asfr} \qquad (5)$$

In the second level (Level 2), the "true EASFRs", unobserved and reconstructed $EASFR^*_{cyae}$, are sampled from a normal distribution centered at the $EASFR^{DHS}_{cyae}$ estimates from the regression model in Step 1 (Equation 1). The standard deviation parameter ν is the standard deviation of the standard errors from the glm estimates and captures to a large extent the variations from the "true EASFRs". We capture standard errors by education level. We allow τ to follow a gamma distribution and produce estimates for each country, year and education level with σ as "initial values" for education level "no education", which was the education level with the highest standard deviation estimate.

Level 2:
$$EASFR^*_{cyae} \sim N(EASFR^{DHS}_{cyae}, \tau_{(EASFR)c,y,e}) \qquad (6)$$
$$\tau_{(EASFR)c,y,e} \sim G(1/\sigma_{(EASFR),e}, 2\sigma^2_{(EASFR)e}) \qquad (7)$$
$$\sigma_{(EASFR),e} = \nu \qquad (8)$$

The third level (Level 3) estimates ESTFRs (education-specific TFR) by sampling from a normal distribution centered on the $ESTFR_{cyae}$ estimates calculated in equation 10 by summing the EASFRs estimated in equation

6 over age groups and multiplying them by 5. The ESTFRs estimates provide a perspective on the evolution of the fertility differentials by level of education at the aggregate level. The standard deviation in this level also follows a normal distribution that centers around the standard deviation from estimates in the regression model. Level 3:

$$ESTFR^*_{cyae} \sim N(ESTFR_{cyae}, \tau_{ESTFR}) \tag{9}$$

$$ESTFR_{cyae} = (\sum_{a=15-19}^{45-49} EASFR^*_{cyae}) \times 5 \tag{10}$$

$$\tau_{ESTFR} = 1/\sigma^2_{ESTFR} \tag{11}$$

$$\sigma_{ESTFR} \sim N(\eta, h) \tag{12}$$

For African countries, ESTFRs are benchmarked to "UN-consistent" ESTFR estimates by Yildiz et. al (2023) which are almost identical to the UN TFRs and thus level 3 is specified as:

$$ESTFR^{un-estfr}_{cyae} \sim N(ESTFR_{cyae}, \tau_{ESTFR}) \tag{13}$$

$$ESTFR^*_{cyae} \sim N(ESTFR_{cyae}, \tau_{ESTFR}) \tag{14}$$

$$ESTFR_{cyae} = (\sum_{a=15}^{49} EASFR^*_{cyae}) \times 5 \tag{15}$$

$$\tau_{(ESTFR)c,y,e} \sim G(\sigma_{(ESTFR),e}, \sigma^2_{(ESTFR)e}) \tag{16}$$

The parameter τ follows a gamma distribution of the standard deviation, σ, of the ESTFR UN-consistent estimates for "higher education".

The full specification of the prior distributions and sensitivity analysis are presented in Appendix B. For Latin American countries, equations (1) to (12) remain and the model for Level 3 does not benchmark to estimates from Yildiz et al. (2023).

Results

In this section, we present estimates of the EASFRs for 1970-1975 to 2015-2020 for all countries under consideration, separately for Africa and Latin America, with a particular emphasis on the starting and end period. Detailed estimates for all five-year periods are presented in the supplementary material.

Weighted EASFRs are calculated for each region in Africa r, education level e, age group a and period y as;

$$\frac{\sum (ASFR_{e,a,y} \times Population_{e,a,y})}{\sum Population_{r,e,a,y}}$$

Figure 3 shows the weighted EASFRs for the African countries under consideration grouped by UN regions. Although the lines between the subsequent levels of education sometimes cross, particularly at older ages - meaning that a higher level of education does not necessarily mean fewer children - each level of education (as categorized in this work) generally leads to lower fertility. In all regions and for all estimated years, women with secondary and/or higher education have the lowest weighted EASFR in their respective age groups, compared to women with primary education or less. Another general observation is that, with increased levels of education, the peak of the fertility curve occurs at older ages; it is particularly true for women with secondary or higher education in comparison with other education groups, throughout the reconstructed period. The low levels of

education of women in Africa translate in the average ASFR curve being very similar in terms of level and pattern with that of the ASFR of women without education in the earlier period, and primary education in the later period. There are notable differences between regions, particularly between Northern and Southern Africa and the other African regions. In the former regions, the ASFR has already converged to the fertility pattern of women with secondary education since the 1990s, indicating greater progress in education in these regions.
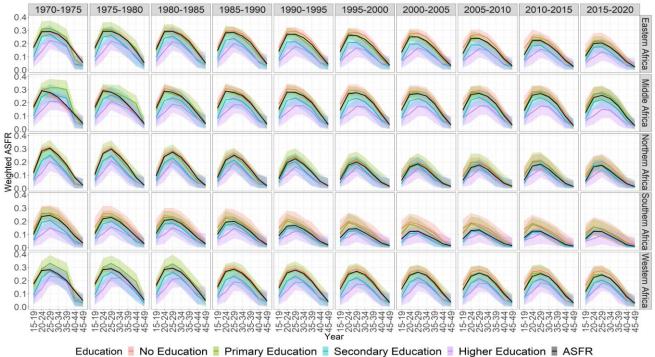
Figure 3: Weighted EASFRs and 95% CI by regions, Africa, 1970-1975 to 2015-2020

Next, we present weighted ESTFRs in Figure 4 estimated for each region in Africa r, education level e, age group a and period y as;

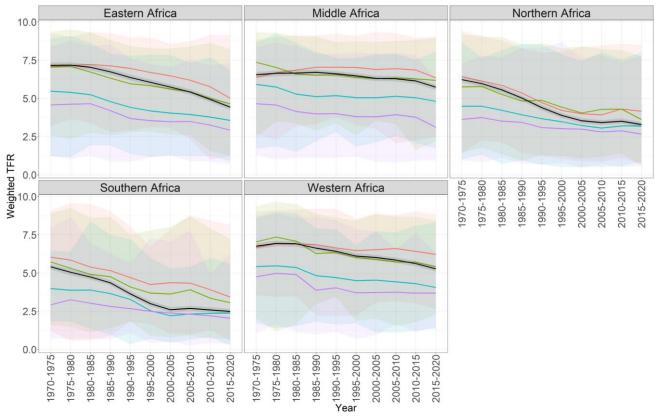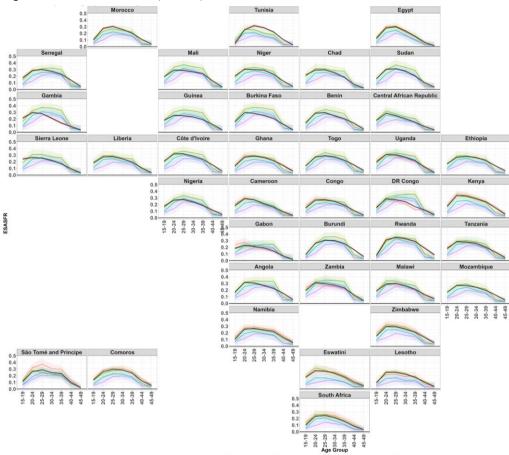$$\frac{\sum \left(TFR_{e,a,y} \times Population_{e,a,y}\right)}{\sum Population_{r,e,a,y}}$$

Figure 4: Weighted ESTFRs and 95% CI by regions, Africa, 1970-1975 to 2015-2020
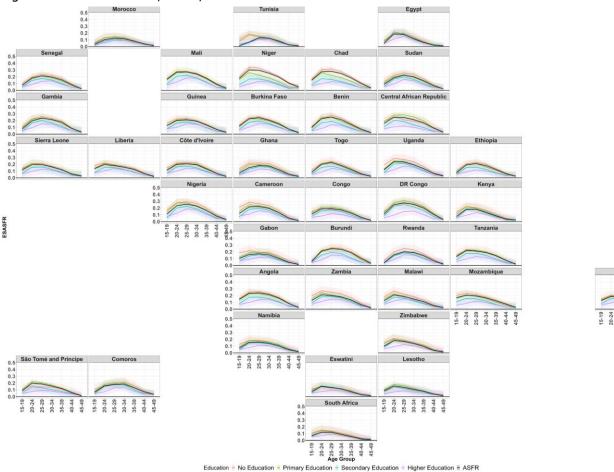

The ESTFR in Middle Africa and Western Africa have been declining very slowly, with some periods of stagnation and even increase, like for women with no education until 2010-2015 in Middle Africa. The stalls in fertility decline have occurred among women in all education categories, such as for women with a primary or a higher education in the 1985-1995 period in Western Africa. In these two regions, the difference in fertility between the women with the highest level of education and the lowest level of education did not decline over time, and has been between 3.6 and 4.9 children per woman between 1970 and 2020. In Eastern Africa, the fertility decline in most education categories is steadier than in the two other regions, especially concerning women with a secondary and higher education. In this region, there seems to be an acceleration of the decreasing trend in the last period since 2010-2015, particularly visible for women without education. This trend is also visible in Western Africa.

Figure 5 and Figure 6 show the EASFR estimates in African countries in the 1970-1975 and 2015-2020 period. Most of the observed patterns in the regions are visible across countries, with the overall ASFRs closely following that of lower education groups.

Figure 5: EASFR and ASFR, Africa, 1970-1975
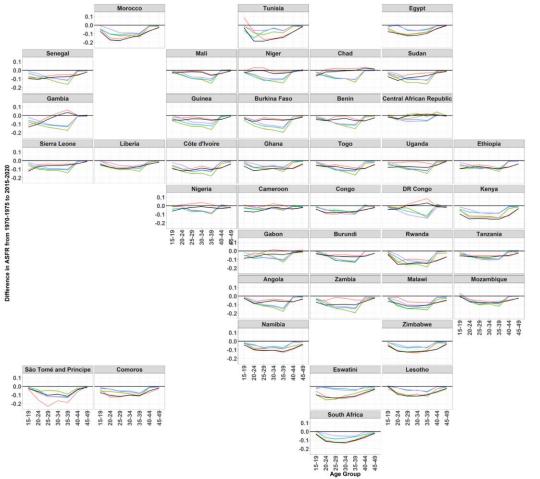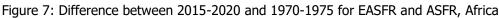
Figure 6: EASFR and ASFR, Africa, 2015-2020

In many African countries, overall ASFR as well as EASFR began to decline significantly from the 1980-1985 period (see supplementary material). On average across all 5-year periods between 1970 and 2020, fertility among women with higher education peaked between the ages of 25 and 34 at around 0.2 to 0.3 children per woman. However, for women with secondary education, fertility was highest overall between the ages of 20 to 29 in all countries, at around 0.2 to 0.4 children per woman. For women with primary or no education, fertility peaked between the ages of 20-25, in general at around 0.3 to 0.4 children per woman.

Figure 7 shows the difference between 2015-2020 and 1970-1975 fertility rates for African countries. For many countries in the region, the difference is negative, implying a decline in fertility rates. The biggest changes are among women with primary education, with a few exceptions. Among all levels of education other than no education, fertility rates have fallen for all ages (excluding ages 45-49) by around 0.1 children per woman, with the exception of Cameroon and Central African Republic.

Figure 7: Difference between 2015-2020 and 1970-1975 for EASFR and ASFR, Africa

Education — No Education — Primary Education — Secondary Education — Higher Education — ASFR

Figure 8 represents the 1970-1975 and 2015-2020 EASFR as well as the difference between the two periods for Latin America. The fertility transition and the education expansion occurred earlier in Latin America compared to Africa (Bongaarts and Casterline 2013; Wils and Goujon 1998). As a result, the differentials in EASFR are larger in 1970-75, and smaller in 2015-2020 when compared with Africa (see Figure 8 panel A and B, and Figure 7). Also, declines in overall ASFR and EASFRs in Latin America are visible from the 1975-1980 period (see supplementary material). The overall ASFR seemed to be largely influenced by the majority of women with primary education in Latin America from the 1970-1975 period until the 1990-1995 period when the overall ASFR became more influenced by the large share of women with a secondary and higher education. This is also reflected in Figure 8 for 1970-1975 and 2015-2020.
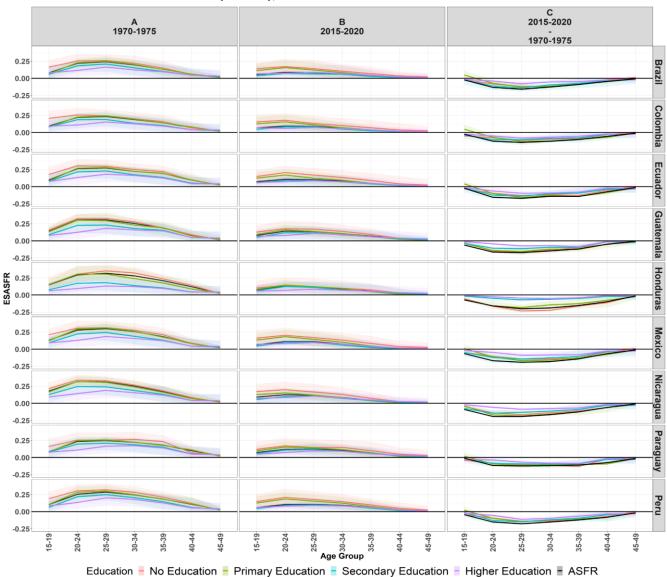
Figure 8: EASFR and ASFR, 1970-1975 (Panel A) and 2015-2020 (Panel B), and difference between 2015-2020 and 1970-1975 for EASFR and ASFR (Panel C), Latin American countries

In all the years under consideration, the fertility peak for women with higher education in Latin America occurred at the age of 25 to 29 years, at a level around 0.2 children per woman in the 1975-1980 period (see supplementary material) and declined to 0.07 children per woman in the 2015-2020 period. For women with secondary education, fertility peaked between the ages of 20 and 29, falling from around 0.25 children per woman in the 1975-1980 period to 0.06 children per woman in the 2010-2015 period. Women with primary education experienced a peak in fertility of around 0.32 children per woman in 1970-1975, declining to 0.14 children per woman in 2000-2005 at the age of 20-24. Similarly, we observed peak fertility rates of about 0.35 children in the 1990-1995 period, falling to 0.14 children in the 2015-2020 period, per woman with no education, at the age of 20-24.

In the 2015-2020 period (Figure 8 panel B), fertility rates in Latin America were significantly lower than in the 1970-1975 period. Panel C shows differences in EASFR between 2015-2020 and 1970-1975. All age groups across all levels of education and countries saw their fertility rates decline. In Brazil, Colombia, Mexico, Ecuador and Peru, women with secondary education between 15-19 and 20-24 saw a large decline of around 0.1 children

per woman. For Peru and Brazil, this decline was reaching almost 0.15 children per woman with a secondary education and was visible for ages until 25-29.

The smallest difference in fertility rates is observed among women with higher education in all the countries studied. In Honduras, women with no education experienced a sharp drop in fertility rates, while in the other countries, it was women with primary education who seemed to experience the largest decrease in fertility rates among all education groups.

Discussion and Conclusion

Our analysis focuses on deriving and applying a methodology to estimate past total and age-specific fertility rates by level of education, for African and Latin American countries. Our aim is to fill a gap by providing a dataset for countries where historical good quality and consistent fertility data are scarce. We propose a Bayesian framework to reconstruct ASFRs by levels of education and provide estimates for 41 African and 9 Latin American countries for five-year periods from 1970 to 2020 using multiple data sources, combining more and less reliable datasets. Our estimates are UN WPP consistent and provide complete age schedules by education levels for past years which were previously not available for Africa and Latin America. However, one of its limitations is that it does not conscientiously model data quality by including the reliability of birth histories that happened a long time before the survey year. Another limitation is the inability to accurately validate the EASFRs, since these are the first estimates of their kind.

Nevertheless, we think these estimates of past fertility rates by age and level of education can be essential in studying in detail the connection to the educational expansion and the role of education in the fertility transition. Furthermore, our estimates can be used to inform population projections. The analysis of the dataset resulting from the modeling confirms what has been demonstrated using existing data, as we will discuss below. These estimates support the general finding that women with higher education have lower fertility rates (e.g., Basu 2002; Bongaarts 2010; Weinberger et al. 1989). The estimates across countries show that women with higher education have a relatively late onset of fertility as well as generally lower fertility rates. The model used can be expanded to other world regions within the DHS database and thus provide reliable estimates for other studies relating to fertility, population, and education.

Many countries in Africa (especially in Sub-Saharan Africa) began their fertility transition in the 1980s (Bongaarts and Casterline 2013). Our estimates of education specific rates in Africa before 1985 agrees with Cochrane (1979) and Martin (1995) that in least developed countries where overall education levels were low, the achievement of some education (incomplete or completed primary education) was sometimes reversely related to fertility. In countries like the Democratic Republic of the Congo, Gabon and Nigeria, fertility rates for women with lower levels of education increased between the 1970-1975 and 2015-2020 periods. The analysis of the stalls in fertility decline show that they could be in part linked to the stalls in education progress (Goujon, Lutz, and KC 2015; Kebede, Goujon, and Lutz 2019). As well, these countries have experienced conflicts in the past, which can be followed by an increase in fertility in the early post-conflict period as shown by the literature (e.g., Lindstrom and Berhanu 1999; Agadjanian and Prata 2002; Randall 2005).

Since Latin America already experienced the fertility transition during the starting years of our analysis, educational differentials in fertility rates were more visible. Our estimates in Latin America for periods before 2000-2005 are in line with the findings of Martin and Juarez (1995) concerning the relatively wide educational fertility differences. However, the differences seem to narrow rather drastically by the end of the period covered by our estimation exercise. The literature points out that the reduction in educational differentials and overall

fertility rates could be attributed to government policies on public health and education, of women in particular (Rios-Neto and Guimarães 2014).

We also note a more pronounced decline in adolescent fertility rates by level of education when considering the differences between 1970-1975 and 2015-2020 in Latin American countries compared to African countries. The decrease in fertility differentials by educational level in many countries in Latin America and some countries in Africa supports the argument of Kravdal (2002), that as more women become educated and reduce their fertility in the community, women with lower levels of education tend to follow this behavior.

That the education expansion has strongly contributed to the decrease in education specific fertility rates between 1970-1975 and 2015-2020 is probable in the countries under consideration. Santelli et al. (2017) report that Latin America and the Caribbean and Sub-Saharan Africa regions spent about 4.4% and 4.6% of their GDP on education in 2012 respectively. This was 42% more than they did in 1990. As a result, the mean years of schooling for women aged 15 to 49 increased on average in Latin America from 4.5 years in 1970 to 9.3 years in 2015, and not as strongly in Africa, which was starting from much lower levels, from 1.3 to 5.9 years (WIC 2018).

References

Abel, G. J. (2021). Wcde: Download data from the wittgenstein human capital data explorer [R package version 0.0.2]. https://CRAN.R-project.org/package=wcde

Agadjanian, V., and Prata, N. (2002). War, peace, and fertility in Angola. Demography, 39(2), 215-231.

Ahuja, V., and Filmer, D. (1995). Educational attainment in developing countries: New estimates and projections disaggregated by gender. Available at SSRN 614964.

Al Zalak, Z., and Goujon, A. (2017). Assessment of the data quality in demographic and health surveys in egypt (tech. rep.). Vienna Institute of Demography Working Papers.

Alkema, L., Raftery, A. E., Gerland, P., Clark, S. J., Pelletier, F., Buettner, T., and Heilig, G. K. (2011). Probabilistic projections of the total fertility rate for all countries. Demography, 48(3), 815–839.

Basu, A. M. (2002). Why does education lead to lower fertility? a critical review of some of the possibilities. World Development, 30(10), 1779–1790.

Bijak, J., and Bryant, J. (2016). Bayesian demography 250 years after bayes. Population studies, 70(1), 1–19.

Bongaarts, J. (2003). Completing the fertility transition in the developing world: The role of educational differences and fertility preferences. Population Studies, 57(3), 321-335.

Bongaarts, J. (2008). Fertility transitions in developing countries: Progress or stagnation? Studies in family planning, 39(2), 105–110.

Bongaarts, J. (2010). The causes of educational differences in fertility in sub-saharan africa. Vienna yearbook of population research, 31–50.

Bongaarts, J., and Casterline, J. (2013). Fertility transition: is sub-Saharan Africa different?. Population and development review, 38(Suppl 1), 153.

Brand, J. E., and Davis, D. (2011). The impact of college education on fertility: Evidence for heterogeneous effects. Demography, 48(3), 863–887.

Bryant, J. R., and Graham, P. J. (2013). Bayesian demographic accounts: Subnational population estimation using multiple data sources. Bayesian Analysis, 8(3), 591-622.

Bumpass, L. L., Rindfuss, R. R., and Jamosik, R. B. (1978). Age and marital status at first birth and the pace of subsequent fertility. Demography, 15(1), 75-86.

Cochrane, S. H. (1979). Fertility and education: World Bank staff occasional papers.

Ellison, J., Dodd, E., and Forster, J. J. (2020). Forecasting of cohort fertility under a hierarchical bayesian approach. Journal of the Royal Statistical Society: Series A (Statistics in Society), 183(3), 829–856.

Fernández-i-Marín, X. (2016). ggmcmc: Analysis of MCMC samples and Bayesian inference. Journal of Statistical Software, 70(9), 1–20. https://doi.org/10.18637/jss.v070.i09

Galor, O. (2012). The demographic transition: Causes and consequences. Cliometrica, 6(1), 1–28.

Gebreselassie, T., and Shapiro, D. (2016). Education and Fertility in Sub-Saharan Africa: What Do We Really Know. ICF International.

Gelman A, Su Y (2022). _arm: Data Analysis Using Regression and Multilevel/Hierarchical Models_. R package version 1.13-1, <https://CRAN.R-project.org/package=arm>.

Goujon, A., Lutz, W., and KC, S. (2015). Education stalls and subsequent stalls in African fertility: A descriptive overview. Demographic Research, 33, 1281-1296.

Hilton, J., Dodd, E., Forster, J. J., Smith, P. W., and Bijak, J. (2019). Forecasting fertility with parametric mixture models. arXiv preprint arXiv:1909.09545.

Jenkins, L., and Macadar, D. (2018). Challenges posed by low fertility in Latin America and the Caribbean.

ICF. 2004-2017. Demographic and Health Surveys (various) [Datasets]. Funded by USAID. Rockville, Maryland: ICF.

Kebede, E., Goujon, A., and Lutz, W. (2019). Stalls in africa's fertility decline partly result from disruptions in female education. Proceedings of the National Academy of Sciences, 116(8), 2891–2896.

Kravdal, Ø. (2002). Education and fertility in sub-Saharan Africa: Individual and community effects. Demography, 39(2), 233-250.

Lee, R. (2003). The demographic transition: Three centuries of fundamental change. Journal of economic perspectives, 17(4), 167–190.

Lesthaeghe, R., Shah, I. H., and Page, H. J. (1981). Compensating changes in intermediate fertility variables and the onset of marital fertility transition. In International Population Conference: Manila (pp. 71-94). Ordina Editions.

Lindstrom, D. P., and Berhanu, B. (1999). The impact of war, famine, and economic decline on marital fertility in Ethiopia. Demography, 36(2), 247-261.

Lutz, W., and Samir, K. (2011). Global human capital: Integrating education and population. Science, 333(6042), 587–592.

Machiyama, K. (2010). A re-examination of recent fertility declines in sub-saharan africa. Marschner, I. C. (2011). glm2: Fitting generalized linear models with convergence problems. The R Journal, 3, 12–15.

Martin, T. C. (1995). Women's education and fertility: results from 26 Demographic and Health Surveys. Studies in family planning, 187-202.

Martin, T. C., and Juarez, F. (1995). The impact of women's education on fertility in latin america: Searching for explanations. International Family Planning Perspectives, 52–80.

Plummer, M. (2021). Rjags: Bayesian graphical models using mcmc [R package version 4-12]. https://CRAN.R-project.org/package=rjags

Pullum, T. W., and Becker, S. (2014). Evidence of omission and displacement in dhs birth histories. ICF International.

Randall, S. (2005). The demographic consequences of conflict, exile and repatriation: A case study of Malian Tuareg. European journal of population/Revue Européenne de Démographie, 21(2), 291-320.

Reher, D. S. (2004). The demographic transition revisited as a global process. Population, space and place, 10(1), 19–41.

Rios-Neto, E. L., and GUIMARÃES, R. D. M. (2014, May). The educational gradient of low fertility in Latin America. In ANNUAL MEETING OF POPULATION ASSOCIATION OF AMERICA. Anais… Boston, MA: PAA.

Santelli, J. S., Song, X., Garbers, S., Sharma, V., and Viner, R. M. (2017). Global trends in adolescent fertility, 1990–2012, in relation to national wealth, income inequalities, and educational expenditures. Journal of adolescent health, 60(2), 161-168.

Schoumaker, B. (2011). Omissions of births in dhs birth histories in sub-saharan africa: Measure- ment and determinants. Population Association of America.

Schoumaker, B. (2013). A stata module for computing fertility rates and tfrs from birth histories: Tfr2. Demographic Research, 28, 1093–1144.

Schoumaker, B. (2014). Quality and consistency of dhs fertility estimates, 1990 to 2012. ICF International Rockville, Maryland.

Schoumaker, B. (2019). Stalls in fertility transitions in sub-saharan africa: Revisiting the evidence. Studies in family planning, 50(3), 257–278.

Ševčiôková, H., Raftery, A. E., and Gerland, P. (2018). Probabilistic projection of subnational total fertility rates. Demographic research, 38, 1843.

Sneeringer, S. (2009). Fertility transition in sub-saharan africa: A comparative analysis of cohort trends in 30 countries (dhs comparative reports no. 23). Calverton, MD: ICF Macro.

Sobotka, T. et al. (2020). Introduction: The relevance of studying fertility across time and space.

Vienna Yearbook of Population Research, 2020.

Speringer, M., Goujon, A., KC, S., Potančoková, M., Reiter, C., Jurasszovich, S., and Eder, J. (2019). Global reconstruction of educational attainment, 1950 to 2015: Methodology and assessment (No. 02/2019). Vienna Institute of Demography Working Papers.

Strulik, H., and Vollmer, S. (2015). The fertility transition around the world. Journal of Population Economics, 28(1), 31–44.

The DHS Program - Available Datasets. (n.d.). https://dhsprogram.com/data/available-datasets.cfm.

United Nations, Department of Economic and Social Affairs, Population Division (2022a). World Population Prospects 2022: Data Sources. (UN DESA/POP/2022/DC/NO. 9).

United Nations, Department of Economic and Social Affairs, Population Division. (2022b). wpp2022: World population prospects 2022 (R package version 1.1-2) [Computer software]. The Comprehensive R Archive Network. Available from http://population.un.org/wpp

United States Agency for International Development. (2013). Standard recode manual for DHS 6.

Weinberger, M. B., Lloyd, C., and Blanc, A. K. (1989). Women's education and fertility: A decade of change in four latin american countries. International Family Planning Perspectives,4–28.

WIC - Wittgenstein Centre for Demography and Global Human Capital. (2018). Wittgenstein Centre Data Explorer Version 2.0. Available at: http://www.wittgensteincentre.org/dataexplorer

Wils, A., and Goujon, A. (1998). Diffusion of education in six world regions, 1960-90. Population and Development Review, 357-368.

Yildiz, D., Wisnioski A.,Brzozowska, Z., and Durowaa-Boateng, A. (2023). A flexible model to reconstruct education-specific fertility rates: Sub-saharan Africa case study. VID Working Paper 2023/02 (forthcoming). Vienna, Austria: Vienna Institute of Demography.

Appendix A - List of Countries and Surveys

Table A1: List of countries and their survey years

Country	Survey Year										
Angola	2006	2011	2015								
Benin	1996	2001	2006	2011	2017						
Brazil	1986	1991	1996								
Burkina Faso	1992	1998	2003	2010	2014	2017					
Burundi	1987	2010	2012	2016							
Cameroon	1991	1998	2004	2011	2018						
Central African Republic	1994										
Chad	1996	2004	2014								
Colombia	1986	1995	2000	2004	2009	2015					
Comoros	1996	2012									
Congo	2005	2011									
Côte D'Ivoire	1994	1998	2011								
DR Congo	2007	2013									
Ecuador	1987										
Egypt	1988	1992	1995	2000	2003	2005	2008	2014			
Eswatini	2006										
Ethiopia	1992	1997	2003	2008							
Gabon	2000	2012									
Gambia	2013	2019									
Ghana	1988	1993	1998	2003	2008	2014	2016				
Guatemala	1987	1995	1998	2014							
Guinea	1999	2005	2012	2018							
Honduras	2005	2011									
Kenya	1988	1993	1998	2003	2008	2014	2015				

Lesotho	2004	2009	2014									
Liberia	1986	2006	2008	2011	2013	2016	2019					
Madagascar	1992	1997	2003	2008	2011	2013	2016					
Malawi	1992	2000	2004	2010	2012	2014	2015					
Mali	1987	1995	2001	2006	2012	2015	2018					
Mexico	1987											
Morocco	1987	1992	2003									
Mozambique	1997	1999	2003	2011	2015							
Namibia	1992	2000	2006	2013								
Nicaragua	1997	2001										
Niger	1992	1998	2006	2012								
Nigeria	1990	2003	2008	2010	2013	2015	2018					
Paraguay	1990											
Peru	1986	1991	1996	2000	2003	2009	2010	2011	2012			
Rwanda	1992	2000	2005	2007	2010	2013	2014	2017				
Sao Tome and Principe	2008											
Senegal	1986	1992	1997	2005	2010	2012	2014	2015	2016	2017	2018	2019
Sierra Leone	2008	2013	2016	2019								
South Africa	1998	2016										
Sudan	1989											
Tanzania	1991	1996	1999	2004	2009	2015	2017					
Togo	1988	1998	2013	2017								
Tunisia	1988	1999										
Uganda	1988	1995	2000	2006	2009	2011	2014	2016				
Zambia	1992	1996	2001	2007	2013	2018						
Zimbabwe	1988	1994	1999	2005	2010	2015						

Appendix B - Prior Distribution and Sensitivity Analysis

We explored different characterization of the parameter ν and how it changed the estimation of EASFRs for both Latin America and Africa models. We keep the model described from equation (2) to (6) and (9) to (12) and change values in equations (7) and (8).

We identify our model described in the text as "Main Model" for both the Latin American and African case. In "Model 1", we define the parameter ν as the standard deviation of the standard error of the glm estimates by country, c, year, y, and age group a and ;

$$\tau_{(EASFR)c,y,a} \sim G(1/\sigma^2_{(EASFR)c,y,a}, \sigma^2_{(EASFR)c,y,a})$$

$$\sigma_{(EASFR)c,y,a} = \nu$$

In "Model 2", we express ν as the standard deviation of the standard error of the glm estimates by education levels only. Then,

$$\tau_{(EASFR)e} \sim G(1/\sigma^2_{(EASFR)e}, \sigma^2_{(EASFR),e})$$

$$\sigma_{(EASFR),e} = \nu$$

In "Model 3", ν is the standard deviation of the standard error of the glm estimates by level of education only. However,

$$\tau_{(EASFR)e} \sim G(1/\alpha_{(EASFR)e}, \sigma^2_{(EASFR),e})$$

$$\sigma_{(EASFR)e} = \nu$$

where $\alpha_{(EASFR)e}$ is the mean of the standard errors of the glm estimates by only level of education.

In "Model 4", we define ν as the standard deviation of the standard errors of the glm estimates. The following are specified for "Model 4",

$$\tau_{EASFR} \sim G(1/\sigma^2_{EASFR}, \sigma^2_{EASFR})$$

$$\sigma_{EASFR} = \nu.$$

We define in "Model 5" the following;

$$\tau_{EASFR} \sim G(1/\sigma^2_{EASFR}, \sigma^2_{EASFR})$$

$$\sigma_{EASFR} = \nu$$

where ν is now the standard deviation of the estimates from the glm model in Step 1.
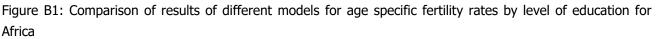
We describe "Model 6" as;

$$\tau_{EASFR} \sim G(\beta_1, \beta_2)$$

and $\beta_1 = (\nu/\alpha)^2$, where ν is the standard deviation of the standard errors from the glm estimates and α is the mean of the standard errors of the glm estimates. Similarly, $\beta_2 = (\nu)^2/\alpha$.

Finally, in "Model 7", we estimate ν as the standard deviation of the standard errors of the glm estimates by each country, year and education level. We specify the following;

$$\tau_{(EASFR)c,y,e} \sim G(1/\sigma^2_{(EASFR)c,y,e}, \sigma^2_{(EASFR)c,y,e})$$

$$\sigma_{(EASFR)c,y,e} = \nu.$$

In Figures B1 and B2, we compare the estimates of "tfr2" module's DHS education specific estimates against the described models to compare how close our models' estimations are to that of DHS.
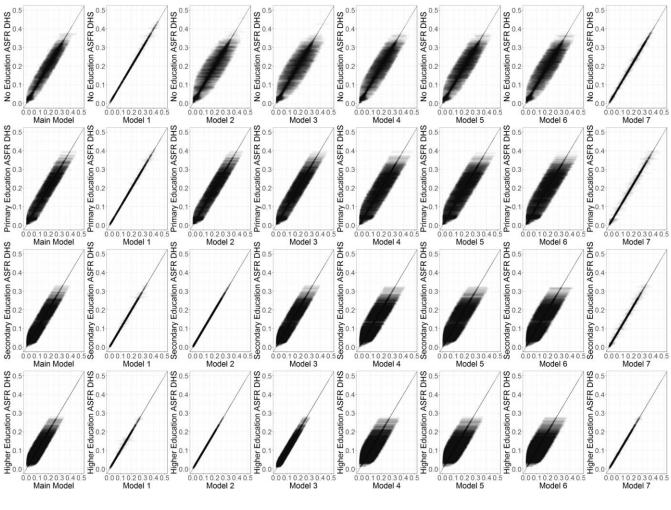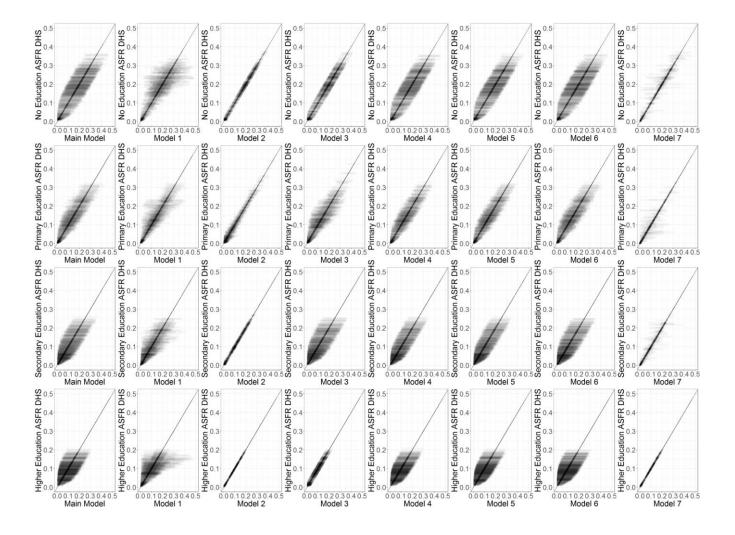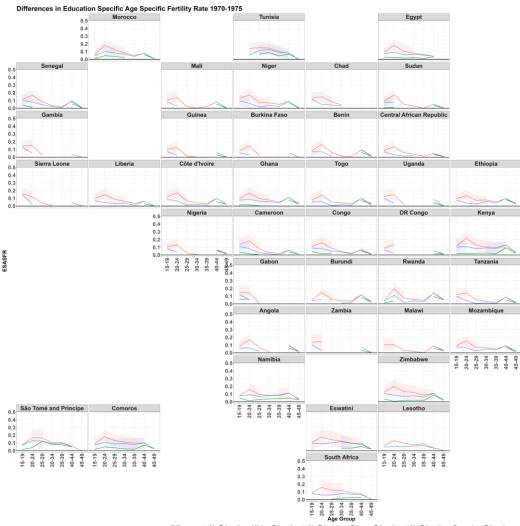
Figure B1: Comparison of results of different models for age specific fertility rates by level of education for Africa

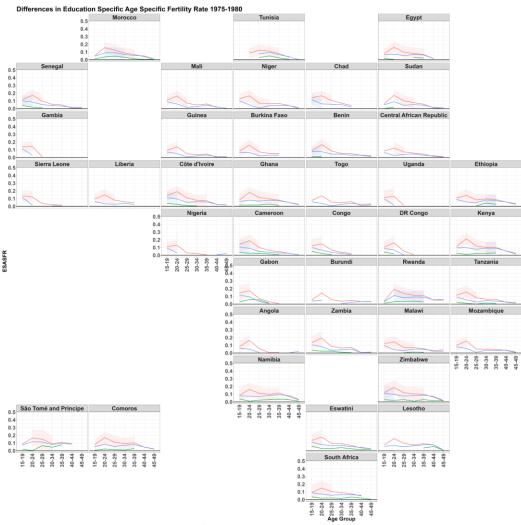Figure B2: Comparison of results of different models for age specific fertility rates by level of education for Latin America

Appendix C - Differences between no education group and other levels of education
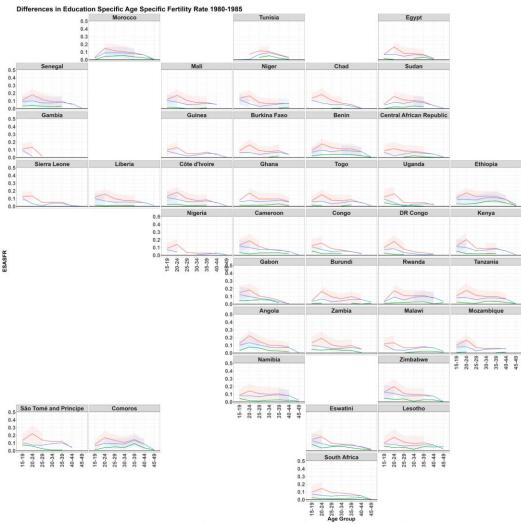

Differences in Education Specific Age Specific Fertility Rate 1970-1975

Differences: — No Education - Higher Education — No Education - Primary Education — No Education - Secondary Education

Differences in Education Specific Age Specific Fertility Rate 1975-1980
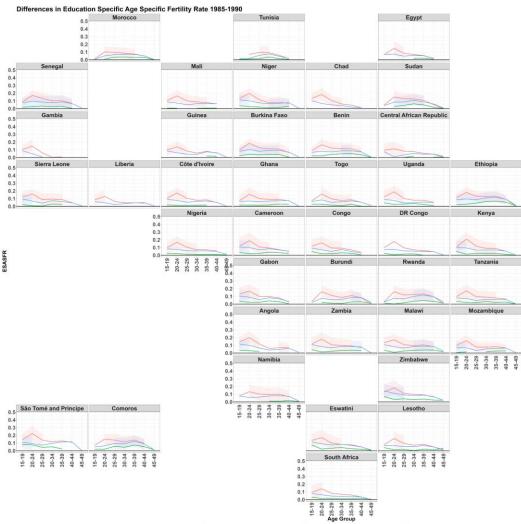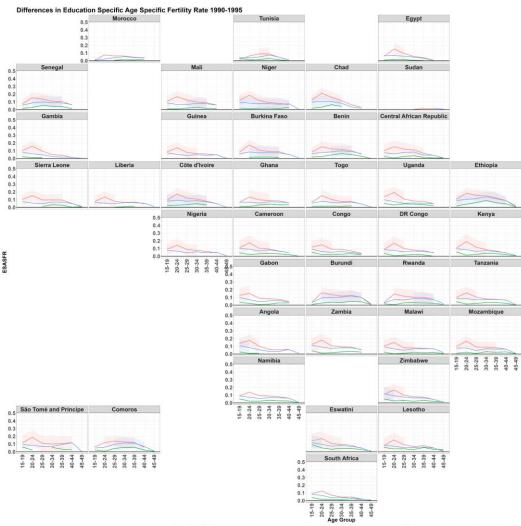
Differences — No Education - Higher Education — No Education - Primary Education — No Education - Secondary Education

Differences in Education Specific Age Specific Fertility Rate 1980-1985

Differences — No Education - Higher Education — No Education - Primary Education — No Education - Secondary Education

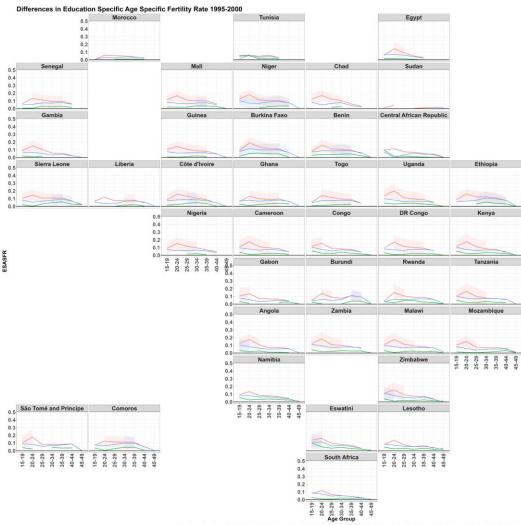Differences in Education Specific Age Specific Fertility Rate 1985-1990

Differences — No Education - Higher Education — No Education - Primary Education — No Education - Secondary Education

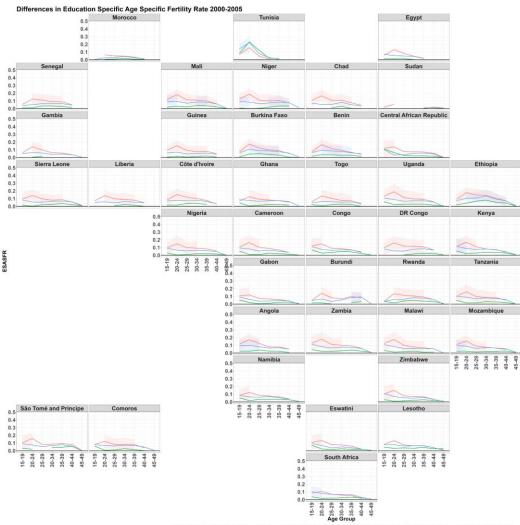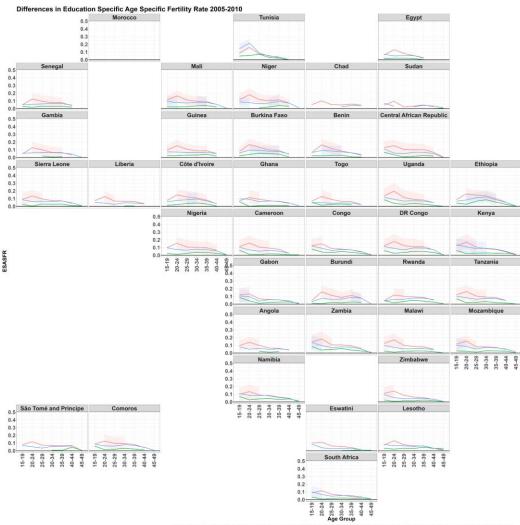Differences in Education Specific Age Specific Fertility Rate 1990-1995

Differences — No Education - Higher Education — No Education - Primary Education — No Education - Secondary Education

Differences in Education Specific Age Specific Fertility Rate 1995-2000

Differences — No Education - Higher Education — No Education - Primary Education — No Education - Secondary Education

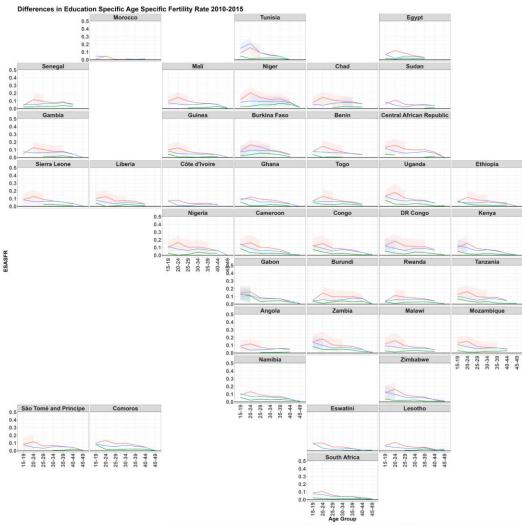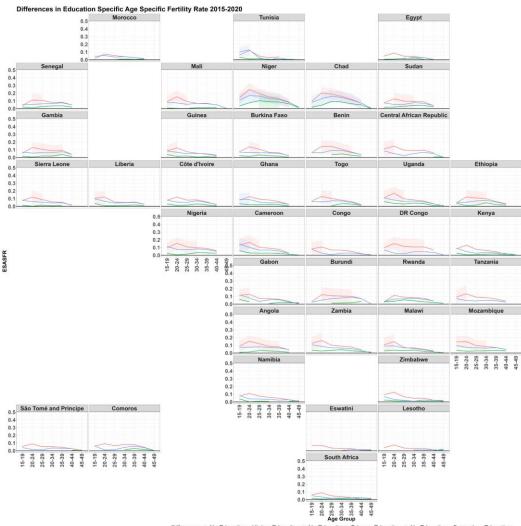**Differences in Education Specific Age Specific Fertility Rate 2000-2005**

Differences — No Education - Higher Education — No Education - Primary Education — No Education - Secondary Education

Differences in Education Specific Age Specific Fertility Rate 2005-2010

Differences — No Education - Higher Education — No Education - Primary Education — No Education - Secondary Education

Differences in Education Specific Age Specific Fertility Rate 2010-2015

Differences — No Education - Higher Education — No Education - Primary Education — No Education - Secondary Education

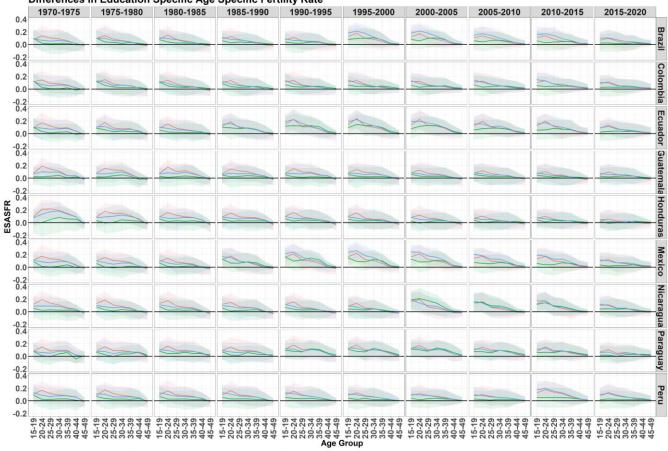Differences in Education Specific Age Specific Fertility Rate 2015-2020

Differences — No Education - Higher Education — No Education - Primary Education — No Education - Secondary Education

Differences in Education Specific Age Specific Fertility Rate

Differences — No Education - Higher Education — No Education - Primary Education — No Education - Secondary Education