Spatial aspects of fertility change in Ethiopia between 2000 and 2016: a district-level analysis

MYUNGGU JUNG

Thesis submitted in accordance with the requirements for the degree of

Doctor of Public Health

of the

University of London

December 2022

Department of Population Health

Faculty of Epidemiology and Population Health

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

Research group affiliation: Population Studies Group (PSG)
No funding was received

Declaration of originality

I, Myunggu Jung, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Myunggu Jung

Signature: [redacted]

Date: 22/12/2022

Abstract

Ethiopia is the second most populous country in sub-Saharan Africa. Although fertility rates in Ethiopia gradually decreased from 5.5 in 2000 to 4.6 in 2016, regional variations in fertility became wider, as evidenced by a TFR of 5.2 in the Somali region and 2.2 in Addis Ababa in 2016.

From the perspective of demographic theory, geographical variations in fertility are often seen as either a reaction to different socioeconomic conditions (the adaptationist approach) or the diffusion of social acceptability of fertility control through geographical distance or linguistic similarity (the diffusionist approach). Recent fertility studies in high- and middle-income countries used spatial models to assess how the adaptation and spatial diffusion effects can jointly account for district-level fertility variations. However, such studies are rare in sub-Saharan Africa due to the shortage of district-level data. This DrPH uses a spatial approach to explore geographical variations in district-level fertility in relation to key selected determinants of fertility for 981 districts using the four Ethiopia Demographic and Health Surveys (2000, 2005, 2011, and 2016).

I began by applying a Bayesian geostatistical approach to estimate the total fertility rate (TFR) and two proximate factors (modern contraceptive prevalence (mCP) and median age at first marriage) and two socioeconomic factors (proportions of women living in urban areas and with secondary education) and one ethnolinguistic factor for 981 districts in 2000, 2005, 2011, and 2016. I found that district-level TFRs within the same region were similar in 2000 and 2005, but they substantially varied in 2011 and 2016. In particular, spatial spreads of lower fertility were observed from the capital city to the northern and western parts of the country in 2011 and 2016.

I then used spatial models to explore spatial autocorrelation of district-TFR and the spatially heterogeneous relationship between TFR and key selected factors affecting TFRs. I found that spatial autocorrelation of TFR became stronger in recent years. Results show that urban-rural differences in fertility were more associated with different socioeconomic conditions, and the recent spatial spread of lower fertility from Addis Ababa to the Amhara region was more associated with spatially heterogeneous effects of mCP, age at marriage, and ethnolinguistic diversity.

This DrPH thesis demonstrates that the geographical location of and distance between districts are important aspects of the recent geographical variations in fertility in Ethiopia. Socioeconomic and cultural characteristics of districts substantially differ even within the same region in Ethiopia, and fertility in a district is affected by where a district is spatially located and the characteristics of nearby districts (diffusion effects), as well as by its own characteristics (adaptation effects). This DrPH thesis provides additional insights into how spatial aspects, as well as socioeconomic, cultural characteristics, and reproductive behaviours in districts, can jointly shape geographical variations in fertility in Ethiopia.

Acknowledgments

I would like to express my gratitude to my three supervisors: Kazuyo Machiyama, Christopher Jarvis, and Ian Timæus. Kazuyo, I consider myself fortunate to have had you as my first supervisor. I am deeply thankful for your invaluable advice and guidance over the past five years, not only in academic matters but also in teaching me how to be a good husband, father, and son. This experience has been truly enlightening in both academic and real-world population studies. I am also grateful for your kindness, patience, and understanding, always accompanied by warm smiles.

Chris, I vividly remember the day I nervously knocked on your office door for the first time, seeking your guidance on my research question and requesting you to be my supervisor. It was undeniably one of the best decisions I made during my DrPH journey. Thank you for helping me navigate the world of R and Spatial analysis, and for your valuable advice on enhancing my writing skills.

Ian, I consider myself privileged to have had you as my supervisor, and I am grateful for the knowledge and insights I gained from you throughout my journey in African demography.

I would also like to express my appreciation to Lynda Clark and Joanna Busza, who supervised my very first 'qualitative' OPA project in Ethiopia. Additionally, I extend my thanks to Gizachew, Taehoon Kim, Heeyeon Kim, and the Population and Youth Section in UNECA for their indispensable support and hospitality, without which my OPA project in Ethiopia would not have been possible.

To my DrPH cohort, I am sincerely grateful for their constant encouragement and motivation. I would also like to acknowledge Prof. Youngtae Cho at Seoul National University, who have consistently broadened my perspective on demography.

Furthermore, I would like to express my gratitude to the organisations where I previously worked, namely COPION, Korea Population Health and Welfare Association (KOPHWA), and Korea Foundation for International Healthcare (KOFIH), for providing me with invaluable experiences in the field of global health and population.

I am deeply grateful to my mother, mother-in-law, brother, and sister for their unwavering love and support. And to my father, I deeply miss you.

Finally, I dedicate this thesis to my beloved Yeohee Choi, who has showered me with love, trust, support, comfort, humour, smiles, and delightful treats such as dumplings, cookies, burgers, lattes, and cheesecake over the past five years. Now, it is my turn. Please give me the opportunity to repay you for the next five thousand years.

Table of contents

List of Figures

List of Tables

Acronyms and Abbreviations

ADP	Amhara Democratic Party
ASFR	Age-specific fertility rates
CI	Confidence intervals
CPR	Contraceptive prevalence rate
DHS	Demographic and Health Surveys
DLDP	District level Decentralisation Programme
DRS	Developing Regional States
EARS	Economically Advanced Regional States
EDHS	Ethiopia Demographic and Health Survey
EFP	European Fertility Project
EPRDF	Ethiopian People’s Revolutionary Democratic Front
GLMM	Generalized linear mixed model
GP	Gaussian Process
GPS	Global Positioning System
GRF	Gaussian random field
GWR	Geographically weighted regression
GTWR	Geographically and temporally weighted regression
HEP	Health Extension Program
HEWs	Health extension workers
HMICs	High- and middle- income countries
HMIS	Health Management Information Systems
HSTP	Health Sector Transformation Plan
INLA	Integrated nested Laplace approximation
LMICs	Low- and middle- income countries
MBG	Model-based geostatistics
MCMC	Markov Chain Monte Carlo
mCP	Modern Contraceptive Prevalence
NPP	National Population Policy
NFFS	National Family and Fertility Survey
UNOCHA	UN Office for Coordination of Humanitarian affairs
ODP	Oromo Democratic Party
OLS	Ordinary least squares
PSUs	Primary sampling units
PPS	Probability Proportional to Size

SAR	Spatial Analysis Reports
SDG	Sustainable Development Goals
SEPDM	Southern Ethiopian People's Democratic Movement
SI	Spatial Interpolation
SLM	Spatial lag model
SNNP	Southern Nation, Nationalities and Peoples
SSA	sub-Saharan Africa
SPDE	Stochastic Partial Differential Equation
TFR	Total fertility rate
TPLF	Tigray People's Liberation Front
WBHSP	Woreda-based Health Sector plan

Notation

A	District (woreda)
a	Age group
β	Regression parameter representing the intervention effect
b	bandwidth size
c	Primary sampling unit (PSU)
$\Gamma(\cdot)$	gamma function
cn	A set of connected neighbouring district
D	Spatial domain
d	Distance
ε	Error term typically $\varepsilon \sim$ N(0, σ2)
g	Number of Ethnic groups
$g(\cdot)$	Link Function
i	Number of location in space
j	Number of district (j = 1, ..., 981)
k	scale parameter
m	Number of explanatory variables
µ	Population mean
N	Number of individuals
n	total number of PSUs
nb	a set of neighbours
ρ	Spatial correlation parameter
s	A location in space
σ	Population standard deviation
T	Matrix transformation
u	Random effect, typically $u \sim$ N(0, σ2)
v	smoothness parameter
W	Spatial weights matrix
X	value of explanatory variable
$Y(\ )$	Outcome variable
z	Number of Zones (z =1…90)
$Z(s)$	Variables at location *s*
$N(\mu, \sigma)$	Gaussian (Normal) distribution
$Bin(N_i, \pi_i)$	Binomial distribution
$GP(0, \Sigma(.))$	Gaussian process with zero mean and covariance functions

<u>Chapter 1</u>

Introduction

1. Chapter 1: Introduction

1.1. Introduction

Interest in understanding local differences in the health and demographics of people grew considerably during the period 2000-2015, sometimes termed the Millennium Development Goal era (WHO, 2016, Hosseinpoor et al., 2018). With this increased focus, it became evident that considering data solely at the country level obscures the important and meaningful differences observed between local areas within a country. Understanding and acknowledging these differences is crucial for effective planning of public health policies and programs, and to ensure that the most vulnerable populations are not overlooked. Consequently, the overarching principle of the post-2015 Sustainable Development Goals (SDG) – "leave no one behind" – emphasizes the need for disaggregated data based on geographic location and other relevant characteristics within national contexts.

Ethiopia, the second most populous country in sub-Saharan Africa (SSA) after Nigeria, is projected to become the eighth largest population in the world by 2050, even with the assumption of a continued decline in fertility rates (UN, 2019). Ethiopia is known for its significant variations in fertility levels across different geographical regions. For instance, in 2016, rural areas had a total fertility rate (TFR) of 5.2, while urban areas had a TFR of 2.2 (ICF, 2016). Despite these variations, very little is known about fertility differences between districts due to the lack of updated district-level data in Ethiopia. Understanding and acknowledging geographical fertility differences between districts are particularly important in Ethiopia, as Ethiopia's districts, *woredas*, are essential administrative units for health policy planning and service delivery. Ethiopia's National Health Sector Transformation Plan 2021-2025, therefore, lays particular emphasis on the '*Woreda transformation*'. Additionally, with more than 80 ethnolinguistic groups residing in Ethiopia, ethnolinguistic identity serves as an important criterion for defining administrative boundaries in the country.

From the perspective of demography, geographical variations in fertility are often seen as either a reaction to different socioeconomic conditions (the adaptationist approach) or the spatial diffusion of new information or social acceptability of fertility control (the diffusionist approach) (Carlsson, 1966). Recent fertility studies in high- and middle-income countries have used spatial models to assess how the adaption and diffusion effects can jointly account for variations in district-level fertility (Campisi et al., 2020, Wang and Chi, 2017, Vitali and Billari, 2017, Sabater and Graham, 2019, Haque et al., 2019). However, such studies have been rare in sub-Saharan African countries. This DrPH thesis contributes to the growing field of research into spatial dimension of fertility to improve our understanding of geographical variations in fertility decline at the district level in sub-Saharan Africa.

1.2. Aim and objectives

The aim of this Doctor of Public Health (DrPH) thesis is to investigate the geographical variations in fertility across 981 districts in Ethiopia from 2000 to 2016, employing geostatistical and spatial modelling approaches. To accomplish this aim, the thesis will address the following research questions and objectives:

<u>Question 1: Are there geographical variations in fertility at the district level between 2000- 2016?</u>

Objective 1 To estimate TFRs and key selected proximate and distal determinants for 981 districts in 2000, 2005, 2011 and 2016 by using a geostatistical modelling approach.

Objective 2 To describe and explore spatial and temporal patterns of TFR and key selected proximate and distal determinants at the district level in 2000, 2005, 2011 and 2016.

<u>Question 2: What determines geographical variations in fertility at the district level?</u>

Objective 3 To assess effects of key selected proximate and distal determinants on geographical variations in fertility at the district level between 2000-2016 with a non-spatial model.

Objective 4 To assess spatial autocorrelation of district-level fertility by using a spatial model.

Objective 5 To explore spatial heterogeneity in relationships between TFRs and both proximate and distal determinants in Ethiopia by using geographically weighted regression between 2000 and 2016.

1.3. Thesis outline

This research paper style thesis contains six chapters. This chapter is an introduction, explaining the aim and objectives. The rest of this thesis has five chapters:

Chapter 2 **Background and theoretical perspective**

Chapter 2 provides backgrounds outlining geographical variations in fertility and provision of population and family planning policies in Ethiopia. Chapter 2 also explains the theoretical perspectives on and determinants of geographical variations in fertility.

Chapter 3 **Data and Methods**

Chapter 3 provides an overview of the data sources and methods used in the two research papers. More detailed information on the methods employed in each of the research papers can be found in Chapters 4 and 5.

Chapter 4 **Are there geographical variations in fertility at the district level between 2000-2016? (Paper 1)**

In Chapter 4, study objectives 1 and 2 are addressed to identify geographical variations in fertility and key selected proximate and distal determinants at the district level in Ethiopia between 2000 and 2016. Model-based geostatistics will be employed to carry out spatial interpolation of fertility and key determinants across the 981 districts. Additionally, spatial and temporal changes in district-level fertility and the key selected determinants will be explored between 2000 and 2016.

Chapter 5 **What determines geographical variations in fertility at the district level? (Paper 2)**

In Chapter 5, I address study objectives 3-5 by using different spatial and non-spatial models to identify and investigate spatial dependency and heterogeneity of district-level fertility in association with the key selected determinants between 2000 and 2016.

Chapter 6 **Discussion and Conclusion**

Chapter 6 presents a comprehensive discussion of the research findings, integrating the results from Chapters 4 and 5. Furthermore, it explores the implications of the findings for population and family planning policies in Ethiopia. Additionally, the chapter provides insights into possible directions for future research.

Each chapter ends with its references.

1.4. References

CAMPISI, N., KULU, H., MIKOLAI, J., KLþSENER, S. & MYRSKYLæ, M. 2020. Spatial variation in fertility across Europe: Patterns and determinants. *Population, Space and Place*, e2308.

CARLSSON, G. 1966. The decline of fertility: innovation or adjustment process. *Population studies,* 20, 149-174.

HAQUE, I., DAS, D. & PATEL, P. P. 2019. Reading the geography of India's district-level fertility differentials: a spatial econometric approach. *Journal of Biosocial Science,* 51, 745-774.

HOSSEINPOOR, A. R., BERGEN, N., SCHLOTHEUBER, A. & GROVE, J. 2018. Measuring health inequalities in the context of sustainable development goals. *Bulletin of the World Health Organization,* 96, 654.

ICF International and Central Statistical Agency. Ethiopia DHS 2016. Addis Ababa, Ethiopia and Calverton, MD: Central Statistical Agency, ICF International, 2016.

SABATER, A. & GRAHAM, E. 2019. International migration and fertility variation in Spain during the economic recession: A spatial Durbin approach. *Applied Spatial Analysis and Policy,* 12, 515-546.

UN 2019. World population prospects 2019. United Nations Geneva.

VITALI, A. & BILLARI, F. C. 2017. Changing determinants of low fertility and diffusion: A spatial analysis for Italy. *Population, Space and Place,* 23, e1998.

WANG, D. & CHI, G. 2017. Different Places, Different Stories: A Study of Spatial Heterogeneity of County-Level Fertility in China. *Demographic research,* 37, 493.

WHO 2016. *World health statistics 2016: monitoring health for the SDGs sustainable development goals,* World Health Organization.

Chapter 2

Background and theoretical perspectives

2. Chapter 2: Background and theoretical perspectives

2.1. Overview

Chapter 2 provides backgrounds outlining geographical variations in fertility and the provision of family planning policies in Ethiopia. This chapter also introduces two theoretical perspectives on geographical variations in fertility and explains the importance of proximate and distal determinants in understanding and explaining geographical variations in fertility in sub-Saharan Africa (SSA).

2.2. Geographical contexts in Ethiopia

2.2.1. Ethnolinguistic geography of Ethiopia

Ethiopia is the most populated landlocked nation in the world and is categorised as a low-income country with a GDP per capita of $890 (2020, current USD). The country shares its borders with Eritrea and Djibouti to the north, Kenya to the south, Somalia to the east, Sudan and South Sudan to the west (Figure 2.1). Ethiopia is primarily a rural and agricultural nation, with approximately 70% of the population engaged in agriculture and 80% of the population residing in rural areas(OECD, 2020).

Ethiopia is the tenth largest country in SSA, but the second most populous country (115 million) in SSA. Ethiopia's population is made up of more than 90 different ethno-linguistic groups and ethno-linguistic identity is the most important criterion for shaping the administrative boundaries between the 11 primary sub-national divisions of the country (Admin 1) (Abbink, 2011, Levine, 2014) (Figure 2.2). This is the result of the 1995 Constitution of Ethiopia based on ethnic-based federalism, dividing Ethiopia along ethnic lines into five regional states dominated by a single ethnic group (Amhara, Tigray, Afar, Oromia, Somiali), four multi-ethnic regional states (Harari, Southern Nation, Nationalities and Peoples (SNNP), Benishangul-Gumuz, Gambella), and two multi-ethnic cities (Addis Ababa and Dire-Dawa) (Erk, 2017, Mengisteab, 1997). The Ethiopian constitution explicitly declares that every ethno-linguistic group has the right to establish self-administrative areas starting at the district (woreda) level, and at zonal and regional levels depending on the size of each ethno-linguistic community (Ethiopia, 1995). This legal framework allows the development and implementation of specific cultural policies within ethno-linguistic boundaries (Abbink, 1998, Abbink, 2011). Consequently, larger ethno-linguistic groups have established their self-administrative areas at the regional level, resulting in regional states being named after their majority ethno-linguistic groups. For example, the Tigray regional state is predominantly inhabited by Tigreans, while the Amhara, Somali, and Oromia regional states are predominantly occupied by Amharas, Somalis, and Oromos, respectively. On the other hand, smaller ethnic groups come together to form "multi-ethnolinguistic" regional states such as SNNP, Gambella, and Benishangul-Gumuz. For instance, the SNNP and Gambella regional states comprise over 50 and 16 indigenous ethno-linguistic groups, respectively. Within these multi-ethnolinguistic regional states, many ethno-linguistic groups have their own administrative zones and districts known as special woredas. For example, the Nuer zone in the Gambela regional state is inhabited by the Nuer ethno-linguistic group, and the Yem special woreda in the SNNP regional state is named after the most populous ethno-linguistic group. However, not all ethno-linguistic groups have their own administrative divisions, and as a result, some newly emphasized ethno-linguistic groups advocate for their separate administrative areas. This is exemplified by the creation of the Sidama and South West regional states in June 2020 and November 2021, respectively, following the Sidama ethno-linguistic group and the people of southwest Ethiopia voting in favour of establishing their own regional states separate from SNNP. Therefore, the formation of the administrative territorial structure based on ethno-linguistic backgrounds is an ongoing process in Ethiopia (Figure 2.1).

This ethnolinguistic-based territorial practice makes Ethiopia unique in SSA, as other SSA countries do not have a formalised system based on ethnolinguistic groups in their geographical and political structure (Abbink, 2011, Abbink, 2009). For example, while ethnicity can play a significant role in the subtext of political elections in many SSA countries, formal ethnic parties are not permitted (Bogaards et al., 2010). In contrast, Ethiopia had an ethnolinguistic federalist political coalition, the Ethiopian People's Revolutionary Democratic Front (EPRDF), leading the government from 1988 to 2019. The EPRDF was a coalition of four ethnolinguistic-based parties: the Oromo Democratic Party (ODP), Amhara Democratic Party (ADP), Tigray People's Liberation Front (TPLF), and Southern Ethiopian People's Democratic Movement (SEPDM). However, in December 2019, the current prime minister, Abiy Ahmed, merged three ethnic-based parties (excluding the TPLF) into the new Prosperity Party. Therefore, until very recently, most political parties in Ethiopia explicitly supported their own ethnolinguistic and regional interests(Østebø et al., 2018).

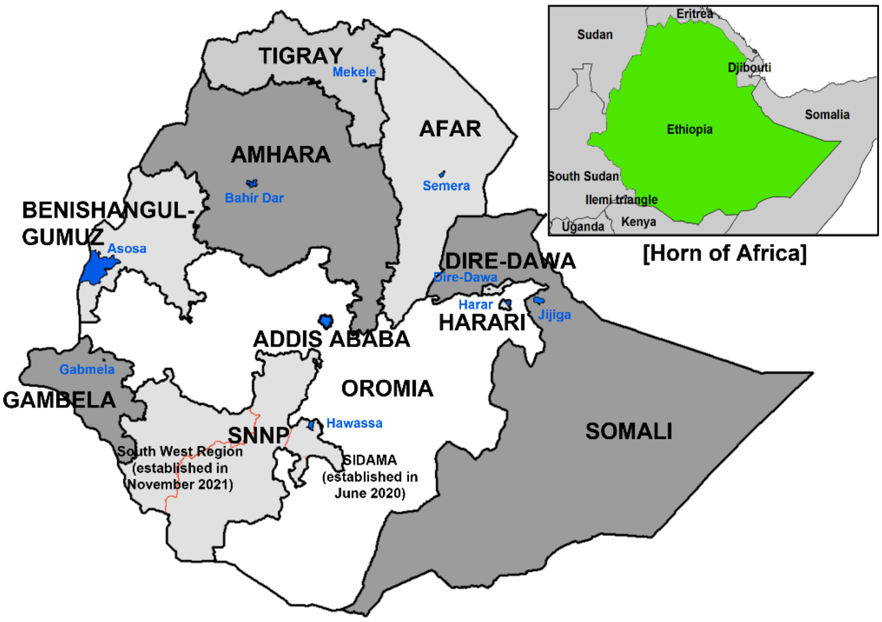


Figure 2.1. Locations of Ethiopia and eleven regional states and regional capitals.
Note: Blue areas on the map indicate the capitals of each region, while red lines depict the boundaries of newly established regions from SNNP. Currently, Hawassa serves as the capital for both the Sidama and the SNNP Regions. However, since Hawassa is located outside of the boundaries of the SNNP region, the SNNP government is planning to identify a new capital and relocate all government institutions within the region's boundaries.

2.2.2. Links between demographic characteristics and geographic areas in Ethiopia

In Ethiopia, it is difficult to track changes in population sizes of different administration areas over time because there have been only three national population and housing census (1984, 1994 and 2007) in Ethiopia. In addition, the 1984 census covered only about 80% of the population and therefore only the 1994 and 2007 censuses covered the entire population. The Ethiopian Statistics Service provides the population projection annually. Although the 2021 population projection indicates that the population is projected to increase by 40% from 73,750,932 in 2007 to 108,998,001 in 2021, the proportions of population in terms of eleven regional states were almost similar between 2007 and 2021(CSA, 2021). Moreover, while the 2021 population projection provides information about population by woredas and urban-rural areas, it does not provide information about population by religion or ethnicity. As a result, this section presents the basic demographic characteristics of Ethiopia by the eleven regional states based on the most recent census conducted in 2007 (Table 2.1).

According to the latest 2007 population and housing census in Ethiopia, the total population of Ethiopia was 73,750,932 (CSA, 2007). Three regional states, namely SNNP, Oromia, and Amhara, accounted for approximately 80% of the total population. The Tigray regional state comprised around 6% of the total population. Additionally, the four regional states of Afar, Gambia, Somalia, and Benishangul-Gumuz are frequently categorised as Developing Regional States (DRS), which are primarily populated by pastoral and migratory communities whose health and socio-economic outcomes lag behind the other regions in Ethiopia. The population of the DRS accounted for 10% of the total population in 2007. Harari regional state, being the smallest regional state in Ethiopia, represented 0.3% of the total population. The two multi-ethnolinguistic cities of Addis Ababa and Dire-Dawa constituted approximately 4.2% of the total population. It is worth noting that the populations of certain woredas, such as Gondar (233,224) in the Amhara region and Adama (220,212) in the Oromia region, exceeded the population of the Harari regional state (183,415).

In terms of religion, the northern parts of Ethiopia, including the Tigray and Amhara regional states, are predominantly inhabited by followers of the Christian Orthodox faith. Conversely, the eastern parts of Ethiopia, including the Somali, Afar, Harari regional states, and Dire Dawa, have a predominantly Muslim population. The relatively ethno-linguistically homogeneous regional states often align with a dominant religion, with the exception of Oromia, which has a diverse religious composition. On the other hand, the multi-ethnic regional states exhibit religious diversity.

In Ethiopia, each ethnic group typically has its own language, and the Ethiopian Constitution also states that "All languages are given equal state recognition" (Article 5.1) and "Every ethnolinguistic group has the right to develop its own language and promote its culture" (Article 39.2). In 2020, the Ethiopian government officially recognised five languages (Amharic, Afaan Oromo, Tigrinya, Somali, and Afar) as the working languages of the federal government. Prior to 2020, Amharic was the only working language of the federal government, and each regional state had the discretion to determine their own working language.

Before 2020, four regional states (Amhara, Benishangul-Gumuz, Gambela, and SNNP) and the two city regions (Addis Ababa and Dire Dawa) chose Amharic as their working language. However, alongside these working languages, there are still numerous minority languages spoken in ethnolinguistic territories. For example, in the SNNP regional state, although Amharic is the working language, students receive 8 years of primary education in their respective native languages.

Despite the distinct territorial differentiation based on ethnolinguistic contexts in Ethiopia, some scholars argue that different regional states, zones, and districts also share common "pan-Ethiopian" themes, such as their history and beliefs in supernatural beings, among other aspects (Levine, 2014, Bach, 2013, Záhořík, 2022). However, it is crucial to acknowledge that the perceived differences between ethnolinguistic groups in Ethiopia are often delineated along geographical boundaries.

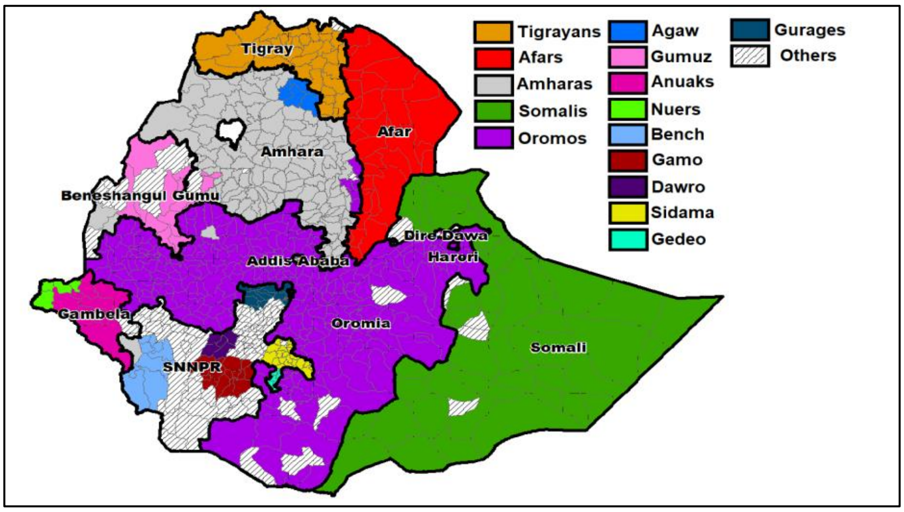


Figure 2.2. Majority ethnic group in each district of Ethiopia according to the 2007 Census
Data source: The 2007 Ethiopia Population and Housing Census

Table 2.1. Ethiopia's demographic characteristics in 11 regional states in 2007

Category	Region	Capital	Population	Residence	Religion	Ethnicity
Urban and Multi-Ethnic region	Addis Ababa	Addis Ababa	2,739,551 (3.7%)	Urban (98.7%) Rural (1.3%)	Orthodox (74.7%) Muslim (16.2%) Protestant(7.77%) Others (1.33%)	Amhara (47.0%) Oromo (19.5%) Gurage (12.3%) Tigrayan(11.2%) Others (10.0%)
	Dire Dawa	Dire Dawa	341,834 (0.5%)	Urban (68.2%) Rural (31.8%)	Muslim (70.8%) Orthodox (25.7%) Protestant (2.8%) Others (0.7%)	Oromo (46.0%) Somali (24.0%) Amhara (20.0%) Gurage (4.5%) Others (5.5%)
	Harari	Harar	183,415 (0.3%)	Urban (54.2%) Rural (45.8%)	Muslim (69.0%) Orthodox (27.1%) Others (3.9%)	Oromo (56.4%) Amhara (22.8%) Harari (8.7%) Gurage (4.3%) Others (7.8%)
Multi-ethnic region	Benishangul-Gumuz (DRS)	Asosa	784,345 (1.1%)	Urban (13.5%) Rural (86.5%)	Muslim (44.9%) Orthodox (33.3%) Protestant(13.5%) Others (8.3%)	Amhara (25.4%) Berta (21.7%) Gumuz (20.9%) Oromo (13.6%) Others (18.4%)
	Gambela (DRS)	Gambela	307,096 (0.4%)	Urban (25.4%) Rural (74.6%)	Protestant(70.1%) Orthodox (16.8%) Muslim (4.9%) Others (8.2%)	Nuer (46.7%) Anuak (21.2%) Amhara (8.4%) Kfficho (5.1%) Others (18.6%)
	SNNP	Hawassa	14,929,548 (20.2%)	Urban (10.0%) Rural (90.0%)	Orthodox (45.6%) Protestant(24.8%) Muslim (16.7%) Others (12.9%)	Gurage(19.5%) Welayta(10.6%) Hadiya(7.9%) Kafficho (5.4%) Others (56.6)
Ethnically homogeneous region	Amhara	Bahir Dar	17,221,976 (23.4%)	Urban (12.3%) Rural (87.7%)	Orthodox (92.5%) Muslim (7.2%) Others (0.3%)	Amhara (91.5%) Others (8.5%)

	Afar (DRS)	Semera	1,390,273 (1.9%)	Urban (13.3%) Rural (86.7%)	Muslim (95.2%) Orthodox (3.9%) Others (0.9%)	Afar (91.8%) Others (8.2%)
	Oromia	Addis Ababa	26,993,933 (36.6%)	Urban (11.3%) Rural (88.7%)	Muslim (47.5%) Orthodox (30.4%) Protestant(17.8%) Others (4.3%)	Oromo (85.0%) Others (15.0%)
	Somali (DRS)	Jijiga	4,445,219 (6.0%)	Urban (20.0%) Rural (80.0%)	Muslim (98.4%) Others (1.6%)	Somalis (99.2%) Others (0.8%)
	Tigray	Mekele	4,316,988 (5.9%)	Urban (19.6%) Rural (80.4%)	Orthodox (95.6%) Muslim (4.0%) Others (0.4%)	Tigrayan (96.6%) Others (3.4%)
Special Enumeration areas			96,754 (0.1%)		-	-
Total			**73,750,932 (100%)**		-	-

* DRS: Developing Regional States
* Special Enumeration areas: areas such as national parks, forest reserves
* Data source: The 2007 Ethiopia Population and Housing Census

2.3. Geographical variations in fertility and family planing policies in Ethiopia

2.3.1. Geographical variations in fertility in Ethiopia

Ethiopia exhibits significant subnational variations in fertility compared to other countries in SSA (Eloundou-Enyegue et al., 2017, ICF, 2016). Ethiopia's fertility levels were on a gradual decline from 2000 to 2016 according to the Ethiopia Demographic and Health Surveys (EDHS) report (Figure 2.3), even though a few studies have pointed out that fertility decline observed in some countries can be spurious due to data quality issues, including age displacement of children age, omissions of children, or different composition of women respondents across the successive surveys (Schoumaker, 2008, Machiyama, 2010). Between 2000 and 2016, the national total fertility rate (TFR) in Ethiopia decreased by 16 percent from 5.52 (95% CI: 5.30-5.74) to 4.56 (95% CI: 4.26-4.87) (ICF, 2000, ICF, 2016). Addis Ababa, for instance, exhibited much lower fertility levels compared to the national average. In 2000, the TFR in Addis Ababa had already declined to 1.8 and further dropped to 1.4 in 2005, which is below the replacement level. Although there was a slight increase to 1.8 in 2016, a recent fertility study across 932 first subnational administrative units (Admin 1) in 70 low-income and middle-income countries, using the most recent DHS data, identified Addis Ababa as having the lowest TFR in SSA (Pezzulo et al., 2021). In terms of trends, the Amhara and SNNP regional states had decreasing trends, while the Afar and Somali regional states experienced increasing trends, albeit some overlap between 95% confidence intervals.

In February 2021, the Ministry of Health of Ethiopia launched the second phase of the Health Sector Transformation Plan (HSTP) II (2020/21-2024/25). The HSTP II recognized the significant variation in Total Fertility Rates (TFRs) across regional states. Therefore, the HSTP II places particular emphasis on addressing equity in healthcare by providing special support to relatively disadvantaged regions (MoH, 2021). In addition to the HSTP II, several studies have explained that regional differentials in fertility in Ethiopia are influenced by multiple factors, such as place of residence (urban vs rural), educational status of women (Shifti et al., 2020, Tessema and Tamirat, 2020), use of contraceptives and access to health facilities (Tessema et al., 2020, Tigabu et al., 2021, Tegegne et al., 2020).

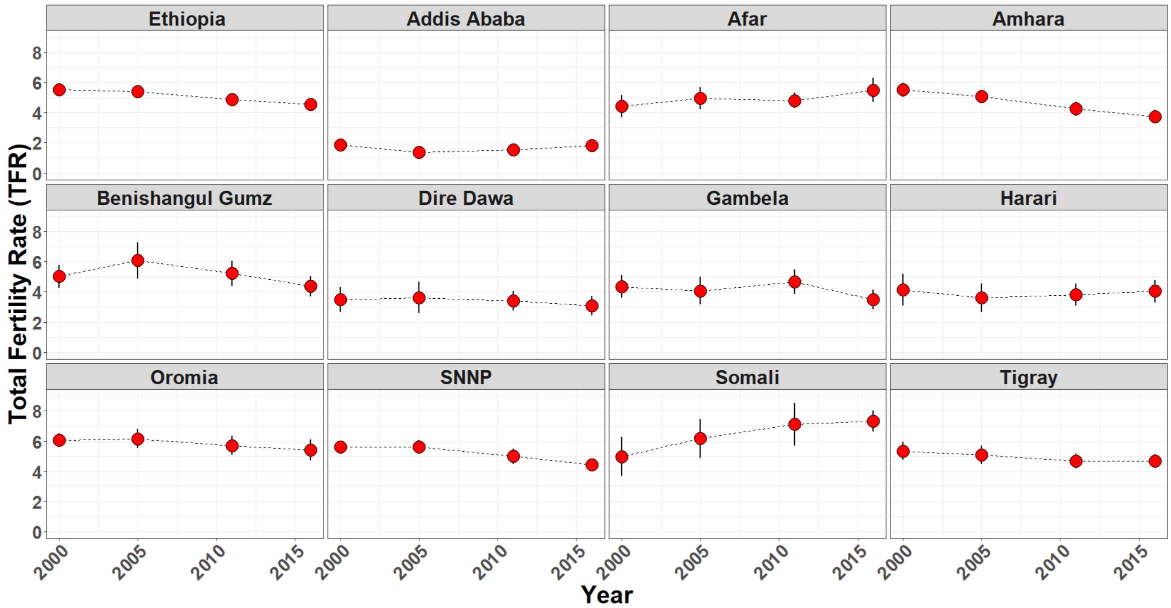


Figure 2.3. National and regional variation of total fertility rates (TFRs) in Ethiopia between 2000 and 2016 with the 95% confidential intervals.
Data sources: Ethiopia Demographic and Health surveys 2000, 2005, 2010 and 2016.

2.3.2. Geographical variation in family planning policy

In 1993, Ethiopia adopted the National Population Policy (NPP), which explicitly acknowledged significant geographical variations in population size between regional states, largely attributed to persistently high fertility levels in certain regions (Ethiopia, 1993). The policy also recognized that high fertility rates could result in regional disparities in socioeconomic and reproductive health outcomes, such as high unemployment and limited access to basic social and reproductive health services. Therefore, a key objective of the Ethiopian population policy was to improve and expand the quality and coverage of family planning services, with the aim of increasing contraceptive prevalence from 4.8%

in 1990 to 44% by the year 2015 (Ethiopia, 1993). To achieve this target, the policy emphasised that a close and functional relationship between the Central government and other regional and local government bodies is vital. Since the implementation of the population policy, there has been a decrease in fertility and an increase in the prevalence of contraception. For instance, total fertility decreased from 5.5 in 2000 to 4.6 in 2016, while the prevalence of contraceptive use among married women rose from 8.1% in 2000 to 35.9% in 2016 (ICF, 2016).

In addition to the national population policy, efforts to reduce regional disparities in reproductive health outcomes, particularly between the DRS (Somali, Afar, Gambella, and Benishangul Gumuz) and other regional states (Tigray, Amhara, Oromia, SNNP, and Harari regional states), were a major focus during the period of the HSTP I (2015/16 – 2019/20) (MoH, 2015). The results of the Ethiopia Demographic and Health Survey (EDHS) showed improvements in some regional states. For instance, in the 2005 EDHS, the prevalence of women using any modern method of contraception was 34.3% in Addis Ababa, 6.6% in Amhara regional state, and 5.0% in SNNP regional state (ICF, 2005). However, in the 2019 Ethiopia Mini DHS, the prevalence increased in all three regional states, and the differences between them became much smaller, with rates of 47.6% for Addis Ababa, 49.5% for Amhara, and 44.6% for SNNP regional states (ICF, 2019).

These achievements were largely attributed to the implementation of the national Health Extension Program (HEP) (Halperin, 2014, May and Rotenberg, 2020). The HEP, initiated in 2004, is a flagship program in Ethiopia that has contributed to improving access to family planning services and played a significant role in the country's recent advancements in reproductive health. The HEP delivers 18 essential health service packages, including family planning, through a workforce of 39,878 health extension workers (HEWs) operating from more than 17,587 health posts in 2019 (MoH, 2020). HEWs actively provide family planning counselling and modern contraception in pastoral, rural, and urban communities. Initially, they offered short-acting methods like condoms, pills, and injectables, but since 2009, their role has expanded to include the authorization to insert implants and provide counselling on where to access removal services (Costenbader et al., 2020). According to the 2019 Mini EDHS, the Amhara regional state reported the largest increase in the use of modern contraceptives between 2005 and 2019, rising from 6.6% to 49.5% (ICF, 2019). A study conducted in 2010 by UNICEF, which involved approximately 400 HEWs and their 10,000 clients, revealed that each health facility had an average of 153 new family planning clients and 157 revisiting clients. The performance in the Amhara regional state was even better, with an average of 245 new clients and 253 revisiting clients per health facility (UNICEF, 2010).

Although Ethiopia has witnessed the remarkable increases in modern contraceptive use, significant regional differences still persist, particularly in DRS, (Lakew et al., 2013, Tegegne et al., 2020, Li et al., 2019a), indicating the need to refocus the family planning program to ensure more equitable access to modern contraceptive methods throughout the country (Shiferaw et al., 2015). In 2000, the modern contraceptive prevalence (mCP) ranged from 2.4% and 7.4% in the Somali and Afar regional states to 8.5% and 12.3% in the Gambela and Benishangul-Gumuz regional states. By 2019, the range had shifted to 3.4% and 12.7% in the Somali and Afar regional states, and 33.2% and 36.7% in the Gambela and Benishangul-Gumuz regional states. Previous studies have identified several reasons for the lower family planning use in the Somali and Afar regional states, including the influence of religious leaders, husbands, and women's attitudes toward family planning methods (Alemayehu et al., 2016, Chekole et al., 2019, Kahsay et al., 2018, Assaf and Wang, 2019, Getnet et al., 2017, Jalu et al., 2019). Moreover, Somali and Afar regional states are home to the largest agro-pastoralist or pastoralist communities in Ethiopia (Teka et al., 2019). These communities are typically nomadic and move seasonally for cattle grazing, making it challenging to access healthcare services, including family planning. Additionally, in agrarian and rural areas, the majority of Health Extension Workers (HEWs) are female, whereas male health workers dominate the pastoralist communities in these two regional states. This suggests that the cultural and environmental factors in Somali and Afar are not conducive to the inclusion and support of females (Getnet et al., 2017). Young (1999) further demonstrated that although Ethiopia's constitution does not make any distinction, a two-tier system of federalism in practice seems to be emerging in Ethiopia (Young, 1999). The highland regional states with higher levels of economic development and political power form one tier, while the four lowland regional states of Gambela, Benishangul-Gumuz, Somali, and Afar, often categorized as the DRS, stand out for their political marginalisation and lower levels of development. Thus, the link between demographic, socioeconomic, and geographical characteristics in the DRS may also contribute to regional differences in health status, including access to modern family planning methods.

Additionally, it should be noted that each region in Ethiopia has a significant geographical size. Consequently, variations in fertility rates and family planning policies are not only expected between regional states but also within them. The HSPT II also calls for innovative solutions to address the health disparities between districts. While several studies in Ethiopia have investigated the determinants of fertility and family planning use in specific zones (Alene and Worku, 2008, Tilahun et al., 2014) and districts (Mekonnen and Worku, 2011, Gebremedhin and Betre, 2009, Atsbaha et al., 2016), there is limited evidence on geographical variations in fertility rates and family planning use across districts in Ethiopia as a whole.

2.4. Rationales for exploring geographical variation in health outcomes at the district level (Admin 3)

To further enhance our understanding of fertility patterns at the district level, it is crucial to recognize the significance of the third-level administrative division, namely the district, in health policy planning and implementation in Ethiopia.

2.4.1. District-based health policy planning and implementation system in Ethiopia

The district-level administrative unit, woreda, holds significant importance for the implementation of population and health policies in Ethiopia. It serves as the third-level administrative division, following the regional states (Admin 1) and zones (Admin 2). While the federal government retains authority over various functions and responsibilities such as fiscal and diplomatic policies, the regional states and woredas have the responsibility of ensuring the provision of basic services within their respective jurisdictions. In particular, District level Decentralisation Programme (DLDP) implemented in 2001/02 played a crucial role in devolving power and resources to woreda governments, while reducing the status and power of zonal administrations (Tesfay, 2015) (Figure 2.4). As a result, woredas have gained the autonomy to develop health policies and deliver services without requiring authorization from zonal governments. Additionally, financial resources are made available through block grant transfers from regional governments to woredas.

The importance of woreda in health policy planning and implementation is prominently emphasised in Ethiopia's National Health Sector Transformation Plan II (HSTP II) 2021-2025. The plan places specific emphasis on the 'Woreda-based Health Sector Plan (WBHSP)', which aims to reduce disparities between high-performing and low-performing woredas (MoH, 2021). The WBHSP is designed to foster alignment and harmonization of health systems for effective planning and monitoring (MoH, 2021).

For instance, in 2017, a malaria elimination program was implemented in 239 targeted woredas across five regional states: Dire Dawa, Harari, Oromia, Tigray, and SNNP (Nega et al., 2020, Assefa et al., 2020). Additionally, since 2017, the Ministry of Health has introduced and promoted the use of a scorecard system at the woreda level to enhance accountability within the health system, with over 600 woredas currently utilizing this system (Argaw et al., 2019). As a result, health policy implementation and targeted interventions are often based on woreda boundaries in Ethiopia. However, despite the significance of understanding and acknowledging the geographical variations in fertility between districts in Ethiopia, there is limited knowledge on fertility differences at the district level due to the lack of updated district-level data in the country.

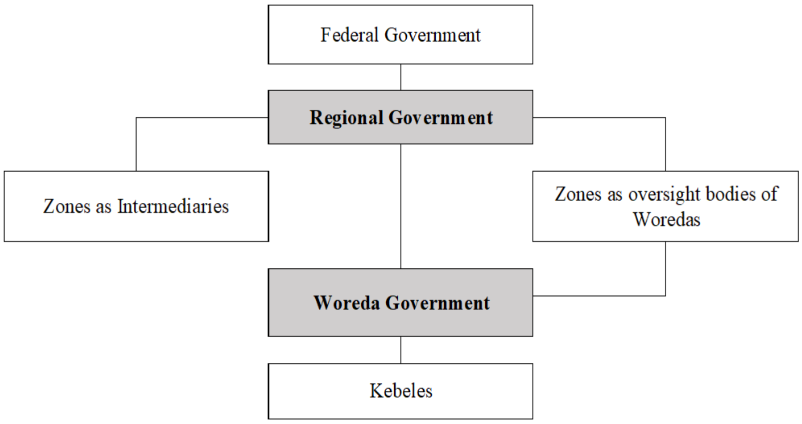


Figure 2.4. Overview of the decentralised governance structure in Ethiopia after the District level Decentralisation Programmes (DLDP)
Source: Zimmermann-Steinhart, P., & Bekele, Y. (2012) (Zimmermann-Steinhart and Bekele, 2012)

2.4.2. Demographic and Health indicator estimates at subnational administrative level

Although the Population and Housing Census (Census) provides essential demographic data at the district level, it is typically conducted every 10 years in most low-income nations, and sometimes even at longer intervals. In the case of Ethiopia, the latest census was conducted in 2007. It is worth noting that the population of Ethiopia is projected to be 102,887,001 in 2021, indicating an increase of approximately 40% from 73,750,932 recorded in the 2007 census, according to the Ethiopian Statistics Service (CSA, 2021). This implies that outdated census data cannot adequately reflect recent variations in demographic and fertility contexts across regions and districts, considering the differential population growth rates and unequal access to family planning services among districts. In addressing the scarcity of updated administrative-level data in low-resource settings, such as many sub-Saharan African countries, Burgert-Brucker et al. (2016)) reviewed three potential approaches (Burgert-Brucker et al., 2016b);

i) **Larger sample sizes**: Increasing sample size for the nationally representative survey to have a representative sample for lower administrative units.
ii) **Routine health management information system**: Using data from community-based healthcare facilities, such as health management information systems (HMIS).
iii) **Model-based geostatistics (MBG) framework:** Producing spatially interpolated maps by using modelling techniques to estimate values at the lower administrative level.

With regard to the first option, it could be possible to carry out a standard DHS that would be representative at the lower administrative levels with larger sample sizes. However, the first option requires substantial financial and human resources to conduct the survey, and therefore, in countries with limited resources, this option may not be practical. Alternatively, a compromise can be made in which some key indicators are available at the lower administrative level (e.g., Admin 2), while others are available only for the higher administrative level (e.g., Admin 1). For instance, the sample size for the 2014 Kanya DHS was about 40,000, and some indicators were available at the second administrative level (47 counties), while others were only available at the first administrative level (8 regions) (ICF, 2014). HMIS data for the second option is often difficult to access, and its data quality is not always reliable. The third option, which makes use of spatial modelling methods, has been increasingly applied for high-resolution mapping of important demographic and health data. The point estimate surface for important demographic and health variables is currently provided by the DHS program in 5 x 5 km pixels and can be downloaded from the Spatial Data Repository (Gething and Burgert-Brucker, 2017).

However, program managers and policymakers typically base their decisions on administrative units rather than the point estimate surface. Therefore, they would be better served by modelled estimates that correspond to their desired administrative units. Recently, the DHS Spatial Interpolation Working Group has also expressed the need for estimates at a lower administrative level than the subnational Admin 1, as health program implementation is decentralized and often occurs at the subnational administrative level 2 (Mayala et al., 2019a, Janocha et al., 2021). The DHS Working Group further demonstrated that Admin 2 estimate maps are useful tools for policymakers, and therefore, these maps should be incorporated into formal decision-making (Kim et al., 2016, Howes et al., 2019, Janocha et al., 2021).

In Ethiopia, the second administrative division (Admin 2) is the Zone. However, as explained earlier, the third-level administrative division (Admin 3), which is the district or woreda, holds more significance for health program implementation and policymaking in Ethiopia. To overcome the shortage of district-level data, the third option is considered to be the least resource-intensive method in Ethiopia for investigating geographical variations in district-level fertility.

2.5. Theoretical backgrounds for geographical fertility variations

There are several theories that attempt to explain geographical variations in fertility. In 1966, Carlsson proposed two approaches known as adaptation and diffusion to explain these variations (Carlsson, 1966). Advocates of the adaptationist approach argue that fertility variations are primarily a response to socioeconomic conditions (Becker, 1960b, Easterlin, 1975). On the other hand, On the other hand, proponents of the diffusionist approach suggest that fertility variations are primarily influenced by the spread of new information or the social acceptability of fertility control methods (Bongaarts and Watkins, 1996, Cleland, 2001, Cleland and Wilson, 1987) .

According to the adaptationist approach, as socioeconomic conditions vary geographically, we would expect variation in area-level socioeconomic conditions to shape geographical fertility differences. However, the adaptationist approach does not always align with observed data. In particular, socioeconomic conditions were found to be only weakly predictive in the Princeton European Fertility Project (EFP) conducted across 1229 provinces and smaller districts in Europe during the 1960s and 1970s (Coale and Watkins, 1986). Hence, the EFP supported the view that it is unlikely that geographical fertility variation can be linked solely to the socioeconomic conditions to which people have adapted.

After the release of evidence from the European Fertility Project, the diffusionist perspective on geographical fertility variation gained momentum. The diffusionist approach originates from the diffusion of innovation theory, which views all social changes through the lens of innovation diffusion. According to the diffusionist approach, the adoption of innovative items is initially slow as it often involves uncertainty and risk. However, the rate of adoption increases quickly due to the social effect of peer groups, as the innovative item becomes more familiar and the element of uncertainty decreases (Rosero-Bixby and Casterline, 1993, Rogers, 1995). Rogers refers to this as the 'diffusion effect.' Therefore, the diffusionist approach argues that geographical fertility variation can be predominantly shaped by the diffusion of new information that influences individuals or communities in their decisions regarding the adoption of deliberate fertility control (Cleland, 2001). The European Fertility Project further demonstrated that fertility changes occurred more rapidly within culturally similar and geographically close populations, irrespective of their high or low socioeconomic conditions (Cleland and Wilson, 1987, Watkins, 1987). Hence, the diffusionist approach claims that socioeconomic development alone is insufficient to account for sub-national fertility differences, and the diffusion of fertility control through different communication pathways should be taken into account (Bongaarts and Watkins, 1996).

Demographers have also shown that the adaptation and diffusion approaches should be viewed as complementary rather than mutually exclusive (Goldstein and Klüsener, 2014, Lesthaeghe and Neels, 2002). For example, individuals become aware of new socioeconomic conditions through both their own socioeconomic circumstances and communication with neighbours (Montgomery and Casterline, 1996). Therefore, uneven development and variations in communication pathways within a country can mutually influence geographical fertility variations.

In addition, historian demographers have stressed the importance of place and context in determining the history of fertility decline (Gillis, Tilly and Levine, 1992). This is particularly because fertility decline is not a unitary national event, and therefore focusing solely on national or regional units to describe fertility decline may mask substantial sub-national differences in fertility decline. Alter (1992) argues that the notion of a unilinear fertility transition, where all societies follow the same fertility decline path, is cast into doubt by data collected for Europe. He particularly pointed out that industrialisation explained only a small portion of the geographical patterns of fertility decline within Europe, and geographical patterns align more closely with linguistic contexts rather than indicators of economic or social development (Alter, 1992). Furthermore, Szreter (1996) proposed the concept of a "*communication community*", where individuals in similar social classes and occupations can have significantly different fertility rates depending on their specific communities. Szreter and his colleagues revealed that the geographical fertility pattern is clearly underlined by the distinct occupational geography of England and Wales, where certain industries came to dominate in particular areas (Garrett, et al., 2001). For instance, in most textile areas in England, female labour force participation rates were high, with opportunities to remain at work after marriage or return to work after having had a child, and marital fertility rates were lower than in most working-class districts (Jaadla et al, 2020). Their works have emphasised that geography and community differences within and between regions both have important roles in determining patterns of fertility. Therefore, local contexts, including shared dialects, norms, and values are crucial elements of contextually varying fertility patterns within a country.

In SSA, Caldwell also wrote that '*most population scholars of the region are guilty of having placed too much emphasis on ... similarities across the [Africa] continent and devoted too little attention to important sub-regional differences*' (Caldwell, 1994a). This is particularly evident in Ethiopia, where exhibits the substantial fertility differences between regional states and the largest urban-rural fertility difference among countries surveyed in the DHS programmes (Appendix 3). As introduced in the section 2.2 and 2.3 in this chapter, Ethiopia's geographical and demographic characteristics such as explicit ethnolinguistic-based geographical boundaries, agro-pastoralist communities concentrated in DRS regions may contribute to these geographical fertility differences. Notably, Ethiopia's capital, Addis Ababa, the capital of Ethiopia, stands out due to its remarkable historical decline in fertility. Although Ethiopia had one of the highest fertility rates of 6.6 among SSA countries in early 1990s, Addis Ababa's fertility rate had declined to nearly the replacement level of 2.2 by 1984 (CSA, 1984) and further dropped below replacement to 1.8 by 1994 (CSA, 1994). Among SSA countries, South Africa has been relatively more industrialized and economically advanced, primarily due to its historical connections with Europe. It is worth noting that the Free State region in South Africa had the lowest fertility rates, reaching 2.2 in 1998 (Department of Health, 2002), which remained higher than the fertility rate observed in Addis Ababa during a similar period. While South Africa exhibited a higher degree of urbanization and had a history of stronger commitment to family planning in SSA during the 1990s (Caldwell, 1994b), this was not the case for Addis Ababa. During the 1990s, the living standards in Addis Ababa, while relatively better than the rest of Ethiopia, remained significantly lower compared to South Africa. Approximately 44% of Addis Ababa's population still lived below 'the poverty line' in 1994 (Tadesse, 1996). Additionally, a study evaluating family planning programmes in low-income countries categorised Ethiopia's program strength in 1994 as 'weak effort' (Ross and Mauldin, 1996). Previous studies have tried to provide explanations for the low fertility rates in Addis Ababa, attributing them to the high economic stress experienced by women of reproductive age (Gurmu and Mace, 2008, Sibanda et al., 2003). They explained that factors such as limited employment opportunities and relatively high housing costs contribute to this economic stress, leading to delayed marriages and reduced marital fertility.

Overall, this approach demonstrates that the sub-national variations in fertility rates highlight certain limitations of fertility transition theory. These variations reveal a picture of multiple, independent fertility declines occurring within a single country, which challenge the assumption of a uniform and linear transition. Therefore, it is essential to acknowledge the significance of geographical place and its local contexts, as well as focusing on the relative influence of economic, cultural, or social factors on a singular process of fertility decline (Gillis, Tilly and Levine, 1992, Boyle, 2003).

2.6. Determinants of geographical variations in fertility

Demographers contend that comprehensive analysis of factors impacting fertility requires that a distinction be made between two sets of determinants: (1) distal and (2) proximate determinants (Bongaarts, 1978). Proximate determinants refer to biological and behavioural factors that directly influence aggregate fertility levels. Distal determinants, on the other hand, encompass socioeconomic, cultural, and other contextual factors that indirectly shape fertility only through their impact on the proximate determinants (Davis and Blake, 1956, Bongaarts and Potter, 1983) (Figure 2.5).

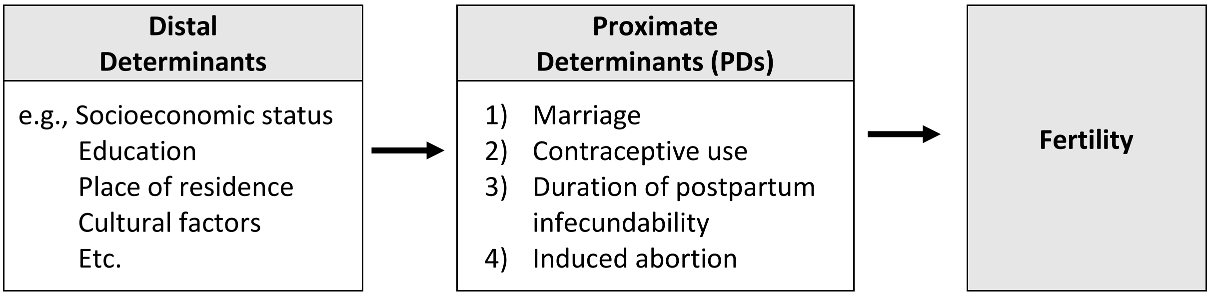


Figure 2.5. Determinants of fertility
Source: (Bongaarts and Potter, 1983)

2.6.1. Distal determinants of geographical variation in fertility

According to Carlsson (1966), if adaptation pressure is the primary driver of geographical fertility patterns, then these patterns would exhibit variations in socioeconomic conditions. On the contrary, if the diffusion effect is the main force shaping these patterns, then they would be structured by variations in communication pathways(Carlsson, 1966).

In terms of variations in socioeconomic conditions, the most commonly identified socioeconomic factors associated with fertility variations are female education levels and urbanization. The adaptationist approach argues that population groups residing in urban areas or with higher levels of education are more likely to engage in the labour market, leading to higher opportunity costs of childbearing (Becker, 1960b). This explanation helps understand the higher fertility levels observed in sub-Saharan African countries, where a lower percentage of the population lives in urban areas and female education levels are generally low (Bryant, 2007, Shapiro and Tenikue, 2017, Behrman, 2015, Kravdal, 2002, Kebede et al., 2019). However, fertility studies in sub-Saharan Africa have shown that the correlations between fertility and the proportion of urban population and education levels are not as strong as observed in non-African low- and middle-income countries (LMICs). This could be attributed to factors such as a higher proportion of people living in urban slums or engaging in informal employment in urban areas (Hassan and Mahabir, 2018, Cáceres-Delpiano, 2012). Furthermore, fertility in African countries has consistently remained higher than in non-African countries due to factors like pronatalist cultural practices(Bongaarts, 2017, Colleran and Snopkowski, 2018, Cleland and Rodriguez, 1988, Caldwell and Caldwell, 1987).

In terms of variations in communication pathways, individuals or communities often vary in their access to information, and the exchange of information is often moderated by cultural and geographical distance (Bongaarts and Watkins, 1996, Hägerstrand, 1965). A well-known study examining the cultural distance in relation to the geographical diffusion of fertility decline is the

Princeton European Fertility Project (EFP), which analysed province-level data in Europe and revealed spatial fertility patterns that closely mirrored Europe's cultural and linguistic geography (Coale and Watkins, 1986). The EFP also demonstrated that fertility decline gradually spread across neighbouring areas, particularly when those areas shared a common language, and the spread appeared to halt at linguistic and cultural boundaries (Watkins, 1987). Geographical fertility studies in local areas of developing countries have also shown that ethnic and cultural diversity can hinder the diffusion of knowledge and attitudes favouring modern reproductive behaviours (De Broe and Hinde, 2006, Yüceşahin and Özgür, 2008, Bongaarts and Watkins, 1996). However, cultural diversification in urban areas can accelerate the diffusion of fertility changes. This is because rural-urban migrants bring with them different cultural practices and ideas, and communication networks in urban areas are more likely to transcend socioeconomic status differences, gender, and ethnicity, facilitating the adoption of innovative reproductive behaviours (Kulu, 2005, Goldstein, 1973, Lee and Farber, 1984, Lerch, 2019, Bongaarts and Watkins, 1996, Klüsener et al., 2019).

In addition, the diffusion of fertility changes can be influenced by geographical distance, as neighbouring areas with similar geographical environments and frequent cross-border commuting may exhibit more similarities in fertility patterns (de Castro, 2007, Goldstein and Klüsener, 2014). This observation aligns with Tobler's first law of geography, which states that "*Everything is related to everything else, but near things are more related than distant things*" (Tobler, 1970). Therefore, when analysing such data, it is important to consider the impact of spatial effects. In spatial statistics, spatial effects are typically categorised into two main terms: spatial dependence and spatial heterogeneity. Spatial dependence in this thesis refers to the degree of the spatial autocorrelation between independently measured values observed in districts, highlighting the tendency for similar values to cluster together in geographical space, regardless of socioeconomic conditions (Kitchin and Thrift, 2009). More specifically, districts in close proximity are inclined to exhibit similar values due to the diffusion process. Spatial heterogeneity in this thesis refers to the variability in relationships across different locations, implying that the association between two study variables can vary instead of being uniform across all places (Anselin, 2010). It recognises that the factors influencing fertility may differ across districts in Ethiopia, leading to diverse patterns of fertility rates at the district level.

Recent studies conducted in high- and middle-income countries (HMICs) using spatial statistical models have identified both spatial dependence and heterogeneity as characteristic features of local fertility patterns. These studies have demonstrated that fertility decline in a particular area is associated with fertility decline in neighbouring areas, even after accounting for socioeconomic and cultural changes (Salvati et al., 2020, Sabater and Graham, 2019, Singh et al., 2017). Moreover, these studies have highlighted significant variations in the relationship between local fertility levels and socioeconomic and cultural factors, both in terms of magnitude and direction (Wang and Chi, 2017, Vitali and Billari, 2017, Campisi et al., 2020, Haque et al., 2019, Costa et al., 2021, Evans and Gray, 2018, Goldstein and Klüsener, 2014). Increasing evidence suggests that the diffusion of new ideas and behaviours may not strictly follow the structure of socioeconomic characteristics. For instance, rapid socioeconomic development in certain areas may not necessarily lead to significant changes in social norms and behaviours. Consequently, some areas may exhibit distinctive fertility behaviour that deviates from the general pattern, while other areas with similar socioeconomic and demographic characteristics may experience different fertility levels (Brunsdon et al., 2002). In Ethiopia, geographical areas often reflect specific ethnolinguistic and cultural attributes within the ethnolinguistic-based territorial structure. Therefore, both empirical evidence and theoretical arguments support the notion that understanding geographical variations in fertility in Ethiopia requires acknowledging the differences in local characteristics. However, studies in sub-Saharan Africa (SSA) have often overlooked the role of spatial effects on geographical variations in fertility, and studies in HMICs have frequently neglected to include proximate variables that could enhance our understanding of the underlying mechanisms driving geographical variations in fertility.

2.6.2. Proximate determinants of geographical variation in fertility

Bongaarts demonstrated that including the proximate variables in the study of the fertility variation is particularly important in SSA, as the offsetting effects of proximate determinants on fertility levels could make relationships between fertility and socioeconomic factors positive, negative or insignificant (Bongaarts et al., 1984). In 1978, Bongaarts proposed a model that incorporates four proximate factors: a) marriage, b) contraception, c) abortion, and d) post-partum infecundability. He provided evidence that the majority of variation in aggregate fertility between geographical regions can be explained by these four factors (Bongaarts, 1978, Bongaarts et al., 1984). Previous studies conducted in SSA countries have utilised the Bongaarts model to examine sub-national variations in fertility and have found that delayed marriage and contraceptive use are two major factors contributing to these variations (Finlay et al., 2018, Singh et al., 1985, Rogers and Stephenson, 2018). Singh et al. specifically noted that postponed marriage often accounts for fertility differences between urban and rural areas, while contraceptive behaviour is a significant factor driving fertility differences among women with different educational levels (Singh et al., 1985). More recent research by Rogers and Stephenson (2018) analysed changes in the proximate determinants of fertility in 82 low- and middle-income countries between 2000-2016. Their findings revealed that the fertility-promoting effects of shorter breastfeeding duration were counterbalanced by increased contraceptive use and delayed marriage, particularly in Eastern and Western African countries(Rogers and Stephenson, 2018). In particular, the study highlighted that the impact of age at marriage on fertility decline was weaker in SSA compared to Latin America and Asia, likely due to the prevalence of premarital births in many SSA countries(Clark et al., 2017). It has been reported that in some SSA countries, such as Botswana and Namibia, more than 60% of first births occur before marriage (Singh, 1998, Garenne and Zwang, 2006). The impact of postponing marriage on fertility can vary depending on the context of premarital childbearing (Harwood-Lejeune, 2001). In settings where premarital childbearing is rare, delaying first marriage can have a greater effect in reducing total fertility levels compared to contexts where premarital childbearing is common. In the Ethiopian context, the country has the lowest median age at first marriage among Eastern African countries, with less than 5% of women giving birth before marriage (Clark and Hamplová, 2013). Consequently, Rogers and Stephenson further revealed that Ethiopia experiences a greater impact of increased age at marriage on fertility levels compared to other SSA nation (Rogers and Stephenson, 2018). Given the similar ages at first marriage and first sexual encounter in Ethiopia (appendix 4), this suggests that the postponement of marriage significantly influences fertility levels in the country.

Sibanda et al. used the 1990 National Family and Fertility Survey (NFFS) and the 2000 Ethiopia Demographic and Health Survey (EDHS) to investigate the decline in fertility in Addis Ababa, which decreased from 3.1 in 1990 to 1.9 in 2000. Their findings indicated that delayed marriage was the most influential proximate determinant contributing to the low fertility level in Addis Ababa (Sibanda et al., 2003). Similarly, Shallo (2020) conducted a study using the EDHS data from 2005, 2011, and 2016 to examine the impact of proximate determinants on fertility decline in Ethiopia. The study concluded that contraceptive use was the most significant proximate determinant accounting for fertility decline over the past decade in Ethiopia (Shallo, 2020). Furthermore, Teklu et al. (2013) analysed the 2000, 2005, and 2011 EDHS data and identified contraception as the primary factor inhibiting fertility among women with secondary and higher educational backgrounds (Teklu et al., 2013).

Hence, contraceptive use has indeed emerged as a crucial proximate determinant of fertility in Ethiopia. The Ethiopian government has implemented effective family planning programs aimed at promoting and facilitating contraceptive uptake. Since 2004, Ethiopia's national family planning program has been implemented through the community-based Health Extension Program (HEP) (Olson and Piller, 2013), which involves the deployment of Health Extension Workers (HEWs) in various local districts, including rural, pastoral, and urban areas (May and Rotenberg, 2020). As a result, variations in contraceptive use between regions and districts are likely to be influenced by different contextual factors such as health infrastructure, ethnolinguistic diversity, and socioeconomic conditions. These variations in contraceptive use at the district level are associated with the geographical variations in fertility levels observed in Ethiopia.

Overall, while the two key proximate determinants of fertility, namely delayed marriage and contraceptive use, play crucial roles in shaping fertility patterns in Ethiopia, there is a limited amount of district-level analysis specifically focusing on proximate determinants in the country. Furthermore, previous studies on district-level fertility in high- and middle-income countries (HMICs) often overlooked the significance of proximate determinants in explaining geographical variations in fertility. As a result, there is a need for more research that examines the district-level dynamics of proximate determinants and their contributions to understanding the geographical variations in fertility in Ethiopia.

2.7. Summary and research gap

There is a clear need for understanding and acknowledging geographical fertility differences between districts in Ethiopia. Furthermore, research gaps in understanding the geographical variation in fertility in Ethiopia can be summarised as follows:

1. Although districts (Admin 3) are essential administrative units for health policy implementation and delivery in Ethiopia, very little is known about geographical variations in fertility at the district level due to the shortage of district-level data in Ethiopia.
2. Although theoretical and empirical studies support the presence of spatial dependency and heterogeneity of fertility in most populations, fertility studies in SSA countries, including Ethiopia, often neglect the spatial effects on sub-national variations in fertility.
3. Recent spatial analysis of fertility in HMICs have revealed the spatially varying relationship between distal determinants and fertility levels at small-scale spatial units. However, these studies often exclude the role of proximate determinants, which are crucial factors accounting for geographical variations in fertility in SSA countries.

Hence, to fill in the gap of research, the aim of this DrPH thesis is to explore geographical variations in fertility in Ethiopia in 981 sub-national areas between 2000 and 2016 by using geostatistical and spatial modelling approaches. This aim will be achieved by addressing the following questions and objectives:

Question 1: Are there geographical variations in fertility at the district level between 2000- 2016?

Objective 1	To estimate TFRs and key selected proximate and distal determinants for 981 districts in 2000, 2005, 2011 and 2016 by using a geostatistical modelling approach.
Objective 2	To describe and explore spatial and temporal patterns of TFR and key selected proximate and distal determinants at the district level in 2000, 2005, 2011 and 2016.

<u>Question 2: What determines geographical variations in fertility at the district level?</u>

Objective 3 To assess effects of key selected proximate and distal determinants on geographical variations in fertility at the district level between 2000-2016 with a non-spatial model.

Objective 4 To assess spatial autocorrelation of district-level fertility by using a spatial model.

Objective 5 To explore spatial heterogeneity in relationships between TFRs and both proximate and distal determinants in Ethiopia by using geographically weighted regression between 2000 and 2016.

2.8. References

ABBINK, J. 2009. The Ethiopian Second Republic and the fragile "social contract". *Africa Spectrum,* 44**,** 3-28.

ABBINK, J. 2011. Ethnic-based federalism and ethnicity in Ethiopia: reassessing the experiment after 20 years. *Journal of Eastern African Studies,* 5**,** 596-618.

ABBINK, J. G. 1998. New configurations of Ethiopian ethnicity: the challenge of the South. *Northeast African Studies,* 5**,** 59-81.

ALEMAYEHU, M., LEMMA, H., ABRHA, K., ADAMA, Y., FISSEHA, G., YEBYO, H., GEBEYE, E., NEGASH, K., YOUSUF, J. & FANTU, T. 2016. Family planning use and associated factors among pastoralist community of afar region, eastern Ethiopia. *BMC women's health,* 16**,** 1-9.

ALENE, G. D. & WORKU, A. 2008. Differentials of fertility in North and South Gondar zones, northwest Ethiopia: A comparative cross-sectional study. *BMC public health,* 8**,** 397.

ALTER, G. 1992. Theories of fertility decline: a non-specialist's guide to the current debate. In J.R. Gillis, L.A. Tilly and D. Levine (eds.), The European Experience of Declining Fertility 1850-1970: The Quiet Revolution. Oxford: Oxford University Press.

ANSELIN, L. 2010. Thirty years of spatial econometrics. Papers in regional science, 89(1), 3-25.

ARGAW, M. D., DESTA, B. F., BELE, T. A. & AYNE, A. D. 2019. Improved performance of district health systems through implementing health center clinical and administrative standards in the Amhara region of Ethiopia. *BMC health services research,* 19**,** 1-13.

ASSAF, S. & WANG, W. 2019. Regional disparities in fertility preferences and demand satisfied for family planning by modern methods across levels of poverty. *DHS Analytical Studies No. 71.* Rockville, Maryland, USA: ICF.

ASSEFA, A., AHMED, A. A., DERESSA, W., WILSON, G. G., KEBEDE, A., MOHAMMED, H., SASSINE, M., HAILE, M., DILU, D. & TEKA, H. 2020. Assessment of subpatent Plasmodium infection in northwestern Ethiopia. *Malaria journal,* 19**,** 1-10.

ATSBAHA, G., HAILU, D., BERHE, H., SLASSIE, A., YEMANE, D. & TEREFE, W. 2016. Determinants of high fertility among ever married women in Enderta district, Tigray region, Northern Ethiopia. *J Health Med Informat,* 7**,** 243-248.

BACH, J.-N. 2013. Compromising with Ethiopianness after 1991: the Ethiopian festival of the millennium (September 2007–September 2008). *Northeast African Studies,* 13**,** 93-122.

BECKER, G. S. 1960b. An Economic Analysis of Fertility, Demographic and economic change in developed countries: a conference of the Universities. *National Bureau Commitee for Economic Research*, 209.

BEHRMAN, J. A. 2015. Does schooling affect women's desired fertility? Evidence from Malawi, Uganda, and Ethiopia. *Demography,* 52**,** 787-809.

BOGAARDS, M., BASEDAU, M. & HARTMANN, C. 2010. Ethnic party bans in Africa: an introduction. *Democratization,* 17**,** 599-617.

BONGAARTS, J. 1978. A framework for analyzing the proximate determinants of fertility.

BONGAARTS, J. 2017. Africa's unique fertility transition. *Population and Development Review,* 43**,** 39-58.

BONGAARTS, J., FRANK, O. & LESTHAEGHE, R. 1984. The proximate determinants of fertility in sub-Saharan Africa. *Population and development review***,** 511-537.

BONGAARTS, J. & POTTER, R. E. 1983. *Fertility, biology, and behavior: An analysis of the proximate determinants,* Academic press.

BONGAARTS, J. & WATKINS, S. C. 1996. Social interactions and contemporary fertility transitions. *Population and development review***,** 639-682.

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

BOYLE, P. 2003. Population geography: does geography matter in fertility research?. *Progress in Human Geography*, 27(5), 615-626.

BRUNSDON, C., FOTHERINGHAM, A. & CHARLTON, M. 2002. Geographically weighted summary statistics—a framework for localised exploratory data analysis. *Computers, Environment and Urban Systems,* 26**,** 501-524.
BRYANT, J. 2007. Theories of fertility decline and the evidence from development indicators. *Population and development review***,** 101-127.
BURGERT-BRUCKER, C. R., DONTAMSETTI, T., MASHALL, A. & GETHING, P. 2016b. Guidance for use of The DHS Program modeled map surfaces. *DHS Spatial Analysis Reports No. 14.* Rockville, Maryland, USA: ICF International.
CãCERES-DELPIANO, J. 2012. Can we still learn something from the relationship between fertility and mother's employment? Evidence from developing countries. *Demography,* 49**,** 151-174.
CALDWELL, J. C. & CALDWELL, P. 1987. The cultural context of high fertility in sub-Saharan Africa. *Population and development review***,** 409-437.
CALDWELL, J. C. 1994a. Fertility in sub-Saharan Africa: status and prospects. P*opulation and Development Review,* 20 (1), 179-193.
CALDWELL, J. C. 1994b. The Course and Causes of Fertility Decline. IUSSP Distinguished Lecture series on Population and Development. LiËge: International Union for the Scientific Study of Population.
CAMPISI, N., KULU, H., MIKOLAI, J., KLþSENER, S. & MYRSKYLæ, M. 2020. Spatial variation in fertility across Europe: Patterns and determinants. *Population, Space and Place***,** e2308.
CARLSSON, G. 1966. The decline of fertility: innovation or adjustment process. *Population studies,* 20**,** 149-174.
CHEKOLE, M. K., KAHSAY, Z. H., MEDHANYIE, A. A., GEBRESLASSIE, M. A. & BEZABH, A. M. 2019. Husbands' involvement in family planning use and its associated factors in pastoralist communities of Afar, Ethiopia. *Reproductive health,* 16**,** 1-7.
CLARK, S. & HAMPLOVã, D. 2013. Single motherhood and child mortality in sub-Saharan Africa: A life course perspective. *Demography,* 50**,** 1521-1549.
CLARK, S., KOSKI, A. & SMITH-GREENAWAY, E. 2017. Recent trends in premarital fertility across sub-Saharan Africa. *Studies in Family Planning,* 48**,** 3-22.
CLELAND, J. 2001. Potatoes and pills: An overview of innovation-diffusion contributions to explanations of fertility decline. *Diffusion processes and fertility transition: Selected perspectives***,** 39-65.
CLELAND, J. & RODRIGUEZ, G. 1988. The effect of parental education on marital fertility in developing countries. *Population studies,* 42**,** 419-442.
CLELAND, J. & WILSON, C. 1987. Demand theories of the fertility transition: An iconoclastic view. *Population studies,* 41**,** 5-30.
COALE, A. J. & WATKINS, S. 1986. The Decline of European Fertility. Princeton: Princeton University Press.
COLLERAN, H. & SNOPKOWSKI, K. 2018. Variation in wealth and educational drivers of fertility decline across 45 countries. *Population Ecology,* 60**,** 155-169.
COSTA, R., BOCQUIER, P. & EGGERICKX, T. 2021. The Contribution of Diffusion to the Fertility Transition in Belgium (1887–1934). *Population and Development Review*.
COSTENBADER, E., CARTWRIGHT, A. F., MCDOWELL, M., ASSEFA, B., TEJEJI, M. Y. & TENAW, E. 2020. Factors Associated With Delayed Contraceptive Implant Removal in Ethiopia. *Global Health: Science and Practice,* 8**,** 000-000.
CSA, 1984. The 1984 population and housing census of Ethiopia. *CSA Addis Ababa*.
CSA, 1994. The 1994 population and housing census of Ethiopia. *CSA Addis Ababa*.
CSA 2007. The 2007 population and housing census of Ethiopia. *CSA Addis Ababa*.
CSA 2021. *Population Projection Wereda as of July 2021,* Addis Ababa, Ethiopian Statistics Service.
DAVIS, K. & BLAKE, J. 1956. Social structure and fertility: An analytic framework. *Economic development and cultural change,* 4**,** 211-235.
DE BROE, S. & HINDE, A. 2006. Diversity in fertility patterns in Guatemala. *Population, Space and Place,* 12**,** 435-459.
DE CASTRO, M. C. 2007. Spatial demography: An opportunity to improve policy making at diverse decision levels. *Population research and policy review,* 26**,** 477-509.

DEPARTMENT OF HEALTH. 2002 South Africa Demographic and Health Survey 1998: Full Report. Pretoria: Department of Health.
EASTERLIN, R. A. 1975. An economic framework for fertility analysis. *Studies in family planning,* 6**,** 54-63.
ELOUNDOU-ENYEGUE, P., GIROUX, S. & TENIKUE, M. 2017. African transitions and fertility inequality: a demographic Kuznets hypothesis. *Population and development review,* 43**,** 59-83.
ERK, J. 2017. 'Nations, Nationalities, and Peoples': The Ethnopolitics of Ethnofederalism in Ethiopia. *Ethnopolitics,* 16**,** 219-231.
ETHIOPIA 1995. *The 1995 Constitution of Ethiopia,* Federal Democratic Republic of Ethiopia, Federal Democratic Republic of Ethiopia.
ETHIOPIA, T. T. G. O. 1993. *National Population Policy of Ethiopia*.
EVANS, A. & GRAY, E. 2018. Modelling variation in fertility rates using geographically weighted regression. *Spatial Demography,* 6**,** 121-140.
FINLAY, J. E., MEJïA-GUEVARA, I. & AKACHI, Y. 2018. Inequality in total fertility rates and the proximate determinants of fertility in 21 sub-Saharan African countries. *PloS one,* 13**,** e0203344.
GARENNE, M. & ZWANG, J. 2006. *Premarital fertility and ethnicity in Africa*, Macro International Calverton.
GEBREMEDHIN, S. & BETRE, M. 2009. Level and differentials of fertility in Awassa town, Southern Ethiopia. *African journal of reproductive health,* 13.
GETHING, P. W. & BURGERT-BRUCKER, C. R. 2017. The DHS Program modeled map surfaces: understanding the utility of Spatial Interpolation for generating indicators at subnational administrative levels. *DHS Spatial Analysis Reports No. 15.* Rockville, Maryland, USA: ICF.
GETNET, F., HASHI, A., MOHAMUD, S., MOWLID, H. & KLINKENBERG, E. 2017. Low contribution of health extension workers in identification of persons with presumptive pulmonary tuberculosis in Ethiopian Somali Region pastoralists. *BMC health services research,* 17**,** 1-9.
GILLIS, J.R, L.A. TILLY and D. LEVINE (eds.). 1992. The European Experience of Declining Fertility 1850-1970: The Quiet Revolution. Oxford: Oxford University Press.
GOLDSTEIN, J. R. & KLþSENER, S. 2014. Spatial analysis of the causes of fertility decline in Prussia. *Population and Development Review,* 40**,** 497-525.
GOLDSTEIN, S. 1973. Interrelations between migration and fertility in Thailand. *Demography,* 10**,** 225-241.
GURMU, E., & MACE, R. 2008. Fertility decline driven by poverty: the case of Addis Ababa, Ethiopia. *Journal of biosocial science*, 40(3), 339-358.
HæGERSTRAND, T. 1965. A Monte Carlo approach to diffusion. *European Journal of Sociology/Archives Européennes de Sociologie/Europäisches Archiv für Soziologie,* 6**,** 43-67.
HALPERIN, D. T. 2014. Scaling up of family planning in low-income countries: lessons from Ethiopia. *The Lancet,* 383**,** 1264-1267.
HAQUE, I., DAS, D. & PATEL, P. P. 2019. Reading the geography of India's district-level fertility differentials: a spatial econometric approach. *Journal of Biosocial Science,* 51**,** 745-774.
HARWOOD-LEJEUNE, A. 2001. Rising age at marriage and fertility in Southern and Eastern Africa. *European Journal of Population/Revue Européenne de Démographie,* 17**,** 261-280.
HASSAN, S. M. & MAHABIR, R. 2018. Urban slums and fertility rate differentials. *Population Review,* 57.
HOWES, R. E., HAWA, K., ANDRIAMAMONJY, V. F., FRANCHARD, T., MIARIMBOLA, R., MIORAMALALA, S. A., RAFAMATANANTSOA, J. F., RAHANTAMALALA, M. A. M., RAJAOBARY, S. H. & RAJAONERA, H. D. G. 2019. A stakeholder workshop about modelled maps of key malaria indicator survey indicators in Madagascar. BioMed Central.
ICF International and Central Statistical Agency. Ethiopia Demogrphic and Health Survey 2005. Addis Ababa, Ethiopia and Calverton, MD: Central Statistical Agency, ICF International, 2005.

ICF International and Kenya National Bureau of Statistics. Kenya Demogrphic and Health Survey 2014. Nairobi, Kenya and Calverton, MD: Kenya National Bureau of Statistics, ICF International, 2014.

ICF International and Central Statistical Agency. Ethiopia Demogrphic and Health Survey 2016. Addis Ababa, Ethiopia and Calverton, MD: Central Statistical Agency, ICF International, 2016.

ICF International and Central Statistical Agency. Ethiopia Demogrphic and Health Survey 2019. Addis Ababa, Ethiopia and Calverton, MD: Central Statistical Agency, ICF International, 2019.

JAADLA, H., REID, A., GARRETT, E., SCHURER, K., & DAY, J. 2020. Revisiting the fertility transition in England and Wales: The role of social class and migration. Demography, 57(4), 1543-1569.

JALU, M. T., AHMED, A., HASHI, A. & TEKILU, A. 2019. Exploring barriers to reproductive, maternal, child and neonatal (RMNCH) health-seeking behaviors in Somali region, Ethiopia. *PLoS one,* 14**,** e0212227.

JANOCHA, B., DONOHUE, R. E., FISH, T. D., MAYALA, B. K. & CROFT, T. N. 2021. Guidance and recommendations for the use of indicator estimates at subnational administrative level 2. *DHS Spatial Analysis Report 20.* Rockville, Maryland, USA: ICF.

KAHSAY, Z. H., TEGEGNE, D., MOHAMMED, E. & KIROS, G. 2018. Application of individual behavioral models to predict willingness to use modern contraceptives among pastoralist women in Afar region, Northern Ethiopia. *PloS one,* 13**,** e0197366.

KEBEDE, E., GOUJON, A. & LUTZ, W. 2019. Stalls in Africa's fertility decline partly result from disruptions in female education. *Proceedings of the National Academy of Sciences,* 116**,** 2891-2896.

KIM, D., SARKER, M. & VYAS, P. 2016. Role of spatial tools in public health policymaking of Bangladesh: opportunities and challenges. *Journal of Health, Population and Nutrition,* 35**,** 1-5.

KITCHIN, R. & TFRIFT, N. 2009. International encyclopedia of human geography. Elsevier.

KLþSENER, S., DRIBE, M. & SCALONE, F. 2019. Spatial and social distance at the onset of the fertility transition: Sweden, 1880–1900. *Demography,* 56**,** 169-199.

KRAVDAL, Ø. 2002. Education and fertility in sub-Saharan Africa: Individual and community effects. *Demography,* 39**,** 233-250.

KULU, H. 2005. Migration and fertility: Competing hypotheses re-examined. *European Journal of Population/Revue européenne de Démographie,* 21**,** 51-87.

LAKEW, Y., REDA, A. A., TAMENE, H., BENEDICT, S. & DERIBE, K. 2013. Geographical variation and factors influencing modern contraceptive use among married women in Ethiopia: evidence from a national population based survey. *Reproductive health,* 10**,** 1-10.

LEE, B. S. & FARBER, S. C. 1984. Fertility adaptation by rural—urban migrants in developing countries: The case of Korea. *Population studies,* 38**,** 141-155.

LERCH, M. 2019. Regional variations in the rural-urban fertility gradient in the global South. *PloS one,* 14**,** e0219624.

LESTHAEGHE, R. & NEELS, K. 2002. From the first to the second demographic transition: an interpretation of the spatial continuity of demographic innovation in France, Belgium and Switzerland. *European Journal of Population/Revue européenne de démographie,* 18**,** 325-360.

LEVINE, D. N. 2014. *Greater Ethiopia: The evolution of a multiethnic society,* University of Chicago Press.

LI, Q., LOUIS, T. A., LIU, L., WANG, C. & TSUI, A. O. 2019a. Subnational estimation of modern contraceptive prevalence in five sub-Saharan African countries: a Bayesian hierarchical approach. *BMC public health,* 19**,** 1-13.

MACHIYAMA, K. 2010. A re-examination of recent fertility declines in Sub-Saharan Africa. *DHS Working Papers No. 68.* Calverton, Maryland, USA: ICF Macro.

MAY, J. F. & ROTENBERG, S. 2020. A Call for Better Integrated Policies to Accelerate the Fertility Decline in Sub-Saharan Africa. *Studies in Family Planning,* 51**,** 193-204.

MAYALA, B. K., DONTAMSETTI, T., FISH, T. D. & CROFT, T. 2019a. *Interpolation of DHS survey data at subnational administrative Level 2,* DHS Program, ICF.

MEKONNEN, W. & WORKU, A. 2011. Determinants of fertility in rural Ethiopia: the case of Butajira Demographic Surveillance System (DSS). *BMC public health,* 11**,** 782.

MENGISTEAB, K. 1997. New approaches to state building in Africa: the case of Ethiopia's ethnic-based federalism. *African Studies Review,* 40**,** 111-132.

MOH 2015. Health sector transformation plan (2015/16–2019/20). Federal Ministry of Health Addis Ababa, Ethiopia.

MOH 2020. *A roadmap for optimizing the Ethiopian Health Extension Program 2020-2035,* Addis Ababa, Ministry of Health-Ethiopia.

MOH 2021. Health Sector Transformation Plan (2020/21-2024/25). Federal Ministry of Health Addis Ababa, Ethiopia.

MONTGOMERY, M. R. & CASTERLINE, J. B. 1996. Social learning, social influence, and new models of fertility. *Population and Development Review,* 22**,** 151-175.

NEGA, D., ABEBE, A., ABERA, A., GIDEY, B., G/TSADIK, A. & TASEW, G. 2020. Comprehensive competency assessment of malaria microscopists and laboratory diagnostic service capacity in districts stratified for malaria elimination in Ethiopia. *PLoS One,* 15**,** e0235151.

OECD 2020. Rural Development Strategy Review of Ethiopia. *OECD*.

OLSON, D. J. & PILLER, A. 2013. Ethiopia: an emerging family planning success story. *Studies in Family Planning,* 44**,** 445-459.

ØSTEBú, M. T., COGBURN, M. D. & MANDANI, A. S. 2018. The silencing of political context in health research in Ethiopia: why it should be a concern. *Health Policy and Planning,* 33**,** 258-270.

PEZZULO, C., NILSEN, K., CARIOLI, A., TEJEDOR-GARAVITO, N., HANSPAL, S. E., HILBER, T., JAMES, W. H., RUKTANONCHAI, C. W., ALEGANA, V. & SORICHETTA, A. 2021. Geographical distribution of fertility rates in 70 low-income, lower-middle-income, and upper-middle-income countries, 2010–16: a subnational analysis of cross-sectional surveys. *The Lancet Global Health,* 9**,** e802-e812.

ROGERS, E. & STEPHENSON, R. 2018. Examining temporal shifts in the proximate determinants of fertility in low-and middle-income countries. *Journal of biosocial science,* 50**,** 551-568.

ROGERS, E. M. 1995. Diffusion of Innovations: modifications of a model for telecommunications. *Die diffusion von innovationen in der telekommunikation.* Springer.

ROSERO-BIXBY, L. & CASTERLINE, J. B. 1993. Modelling diffusion effects in fertility transition. *Population studies,* 47**,** 147-167.

ROSS, A.J. and MAULDIN, P. 1996. Family planning programs: efforts and results, 1972-94. *Studies in Family Planning,* 27 (3): 137-147.

SABATER, A. & GRAHAM, E. 2019. International migration and fertility variation in Spain during the economic recession: A spatial Durbin approach. *Applied Spatial Analysis and Policy,* 12**,** 515-546.

SALVATI, L., BENASSI, F., MICCOLI, S., RABIEI-DASTJERDI, H. & MATTHEWS, S. A. 2020. Spatial variability of total fertility rate and crude birth rate in a low-fertility country: patterns and trends in regional and local scale heterogeneity across Italy, 2002–2018. *Applied Geography,* 124**,** 102321.

SCHOUMAKER, B. Stalls and reversals in fertility transitions in sub-Saharan Africa: Real or spurious? Human Fertility in Africa: Trends in the last decade and prospects for change, 2008.

SHALLO, S. A. 2020. Roles of proximate determinants of fertility in recent fertility decline in Ethiopia: Application of the revised bongaarts model. *Open Access Journal of Contraception,* 11**,** 33.

SHAPIRO, D. & TENIKUE, M. 2017. Women’s education, infant and child mortality, and fertility decline in urban and rural sub-Saharan Africa. *Demographic Research,* 37**,** 669-708.

SHIFERAW, S., ABDULLAH, M., MEKONNEN, Y., MAñGA, A., AKINYEMI, A., AMOUZOU, A., FRIEDMAN, H., BARROS, A. J. & HOUNTON, S. 2015. Trends in contraceptive use

and distribution of births with demographic risk factors in Ethiopia: a sub-national analysis. *Global health action,* 8**,** 29720.
SHIFTI, D. M., CHOJENTA, C., HOLLIDAY, E. G. & LOXTON, D. 2020. Application of geographically weighted regression analysis to assess predictors of short birth interval hot spots in Ethiopia. *PloS one,* 15**,** e0233790.
SIBANDA, A., WOUBALEM, Z., HOGAN, D. P. & LINDSTROM, D. P. 2003. The proximate determinants of the decline to below-replacement fertility in Addis Ababa, Ethiopia. *Studies in family planning,* 34**,** 1-7.
SINGH, A., KUMAR, K., PATHAK, P. K., CHAUHAN, R. K. & BANERJEE, A. 2017. Spatial patterns and determinants of fertility in India. *Population,* 72**,** 505-526.
SINGH, S. 1998. Adolescent childbearing in developing countries: a global review. *Studies in family planning***,** 117-136.
SINGH, S., CASTERLINE, J. B. & CLELAND, J. G. 1985. The proximate determinants of fertility: Sub-national variations. *Population studies,* 39**,** 113-135.
SZRETER, S. 1996. Fertility, class and gender in Britain, 1860-1940. Cambridge: Cambridge University Press.
TADESSE, M. 1996. Food consumption and poverty in urban Ethiopia: a preliminary assessment. Pp. 215-225 in B.Kebede and M. Tadesse (eds.), The Ethiopian Economy: Poverty and Poverty Alleviation, Proceedings of the Fifth Annual Conference on the Ethiopian Economy. Addis Ababa: Department of Economics, Addis Ababa University and Ethiopian Economic Association.
TEGEGNE, T. K., CHOJENTA, C., FORDER, P. M., GETACHEW, T., SMITH, R. & LOXTON, D. 2020. Spatial variations and associated factors of modern contraceptive use in Ethiopia: a spatial and multilevel analysis. *BMJ open,* 10**,** e037532.
TEKA, A. M., WOLDU, G. T. & FRE, Z. 2019. Status and determinants of poverty and income inequality in pastoral and agro-pastoral communities: Household-based evidence from Afar Regional State, Ethiopia. *World Development Perspectives,* 15**,** 100123.
TEKLU, H., SEBHATU, A. & GEBRESELASSIE, T. 2013. *Components of Fertility Change in Ethiopia: Further Analysis of the 2000, 2005, and 2011 Demographic and Health Surveys*, ICF International.
TESFAY, A. 2015. *The extent and impacts of decentralization reforms in Ethiopia.* Boston University.
TESSEMA, Z. T., AZANAW, M. M., BUKAYAW, Y. A. & GELAYE, K. A. 2020. Geographical variation in determinants of high-risk fertility behavior among reproductive age women in Ethiopia using the 2016 demographic and health survey: a geographically weighted regression analysis. *Archives of Public Health,* 78**,** 1-12.
TESSEMA, Z. T. & TAMIRAT, K. S. 2020. Determinants of high-risk fertility behavior among reproductive-age women in Ethiopia using the recent Ethiopian Demographic Health Survey: a multilevel analysis. *Tropical Medicine and Health,* 48**,** 1-9.
TIGABU, S., LIYEW, A. M. & GEREMEW, B. M. 2021. Modeling spatial determinates of teenage pregnancy in Ethiopia; geographically weighted regression. *BMC Women's Health,* 21**,** 1-12.
TILAHUN, T., COENE, G., TEMMERMAN, M. & DEGOMME, O. 2014. Spousal discordance on fertility preference and its effect on contraceptive practice among married couples in Jimma zone, Ethiopia. *Reproductive health,* 11**,** 1-10.
TOBLER, W. R. 1970. A computer movie simulating urban growth in the Detroit region. *Economic geography,* 46**,** 234-240.
UNICEF 2010. *HEWs and Health Post Performance and Community Satisfaction and Perception: HEP Evaluation in Rural Ethiopia,* Addis Ababa, UNICEF.
VITALI, A. & BILLARI, F. C. 2017. Changing determinants of low fertility and diffusion: A spatial analysis for Italy. *Population, Space and Place,* 23**,** e1998.
WANG, D. & CHI, G. 2017. Different Places, Different Stories: A Study of Spatial Heterogeneity of County-Level Fertility in China. *Demographic research,* 37**,** 493.
WATKINS, S. C. The fertility transition: Europe and the Third World compared. Sociological forum, 1987. Springer, 645-673.

YOUNG, J. 1999. Along Ethiopia's western frontier: Gambella and Benishangul in transition. *The journal of Modern African studies,* 37**,** 321-346.
YþCEŞAHIN, M. M. & ÖZGþR, E. M. 2008. Regional fertility differences in Turkey: persistent high fertility in the southeast. *Population, Space and Place,* 14**,** 135-158.
ZãHOŘïK, J. 2022. Competing-Isms in the Horn of Africa: The Rise and Fall of Pan-Ethiopianism and Pan-Somalism. *Nationalism and Ethnic Politics,* 28**,** 74-91.
ZIMMERMANN-STEINHART, P. & BEKELE, Y. 2012. The implications of federalism and decentralisation on socio-economic conditions in Ethiopia. *Potchefstroom Electronic Law Journal/Potchefstroomse Elektroniese Regsblad,* 15.

Chapter 3

Data and Methods

3. Chapter 3: Data and Methods

3.1. Overview

This section introduces the data and methods for each paper and study objective in this thesis, as summarized in Table 3.1. Paper 1 focuses on estimating total fertility rates (TFR) and key selected determinants of fertility in 981 districts to describe changes in geographical variations in each study variable between 2000 and 2016. Building upon Paper 1, Paper 2 further explores the spatial relationship between TFR and the selected determinants at the district level.

Table 3.1. Summary of the data and study methods

	Study Objective	Data	Study method
Paper 1: Are there geographical variations in fertility at the district level in Ethiopia?			
Chapter4	**Objective 1:** To estimate total fertility rates and key selected proximate and distal determinants for 981 districts in 2000, 2005, 2011 and 2016 by using a geostatistical modelling approach.	**Secondary Data:** 1. Ethiopia Demography and Health Surveys 2000, 2005, 2011, 2016 2. Ethiopian Statistics Service-Subnational Administrative Boundaries (shapefiles)	Model-based geostatistics with INLA and SPDE
	Objective 2: To describe and explore spatial and temporal patterns of fertility and key selected determinants at the district level between 2000 - 2016	**Secondary Data:** 1. Fertility & related-indicator estimates at 981 districts * 2. Subnational Administrative Boundaries (shapefiles)	Mapping of fertility and key indicator estimates at 981 districts
Paper 2: What determines geographical variations in fertility at the district level?			
Chapter5	**Objective 3:** To assess effects of key selected distal and proximate determinants on geographical fertility variation at the district level between 2000-2016 with non-spatial linear model	**Secondary Data:** 1. Ethiopia Demography and Health Surveys 2000, 2005, 2011, 2016 2. Fertility & related-indicator estimates at 981 districts * 3. Subnational Administrative Boundaries (shapefile)	- Non-spatial linear regression model - Semi-variogram
	Objective 4: To assess spatial dependency of local fertility by using spatial linear model	**Secondary Data:** 1. Fertility & related-indicator estimates at 981 districts * 2. Subnational Administrative Boundaries (shapefile)	- Spatial Linear regression model: (Spatial Lag Model)
	Objective 5: To explore spatial heterogeneity in relationships between district-level fertility and both proximate and distal determinants in Ethiopia by using GWR model between 2000 and 2016	**Secondary Data:** 1. Fertility & related-indicator estimates at 981 districts * 2. Subnational Administrative Boundaries (shapefile)	- Geographically Weighted Regression (GWR) model

* Modelled data generated from the outcome of Objective 1

3.2. Paper 1: Are there geographical variations in fertility at the district level in Ethiopia?

For paper 1, I use 2000, 2005, 2011, and 2016 Ethiopia Demographic and Health Surveys (EDHS) to predict TFRs and key selected determinants at 981 districts in Ethiopia by using a model-based geostatistical approach. In addition, I use the district-level fertility and key selected indicator estimates to describe changes in spatial and temporal patterns of district-level fertility and key selected determinants between 2000 and 2016. This research paper addresses study objectives 1 and 2 as outlined in Table 3 (Table 3.1).

3.2.1. Objective 1: To describe and explore spatial and temporal patterns of study variables in 2000, 2005, 2011 and 2016

3.2.1.1. Data

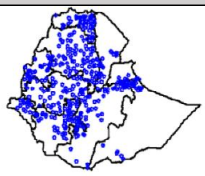
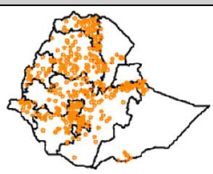
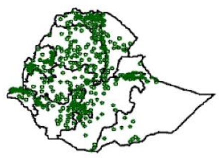
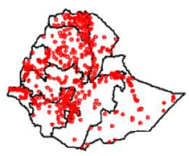
I obtained data from Ethiopia Demographic and Health Surveys (EDHS) conducted in 2000, 2005, 2011, 2016 (ICF, 2000, ICF, 2005, ICF, 2011, ICF, 2016). I used two data sets from each EDHS; the Individual Women's Recode data and the GPS data. The Individual Women's Recode data contains information about backgrounds, reproductive health of women of reproductive age (15-49) and a full history listing all live births that female respondents have given birth to. The GPS data provides the latitude, longitude for sampling units, and therefore the georeferenced datasets can be linked to the Individual Women's Recode dataset through unique identifiers of each sampling unit.

The EDHS samples are designed to give indicator estimates that are nationally representative, as well as representative for the 11 regional states. The four EDHS samples (2000, 2005, 2011 and 2016) are stratified by regional states and urban/rural areas and the samples were selected in two stages. In the first stage of sampling, census enumeration areas (EAs) are selected in the first stage as primary sampling units (PSUs) with probability proportional to size (PPS). In the second stage, approximately 24-35 residential households are randomly selected from each selected PSU after a household listing operation has generated an updated, complete list of residential households in the selected PSU. To ensure that households and respondents' confidentiality is maintained, especially in clusters with a small number of sampled households and respondents, the DHS apply a random displacement to the GPS latitude/longitude coordinates before releasing data externally. Therefore, they are an estimated centre of a cluster of households, which are point locations that actually represents an area of unknown size with fairly large variability across a country, especially between urban and rural locations. In addition, these point locations are geo-masked from 0-2 km in urban locations and 0-5 km in rural locations, with 1% of rural locations up to 10 km (Perez-Haydrich et al., 2013, Burgert et al., 2013).

Of the total number of PSUs, I excluded 6, 9, 27, and 21 PSUs without GPS codes, resulting in the inclusion of 533, 526, 569, and 622 PSUs for the 2000, 2005, 2011, and 2016 EDHS, respectively (Table 3.2). It is important to note that due to security concerns in the Afar and Somali regions, a small number of zones were included in the 2000 and 2005 EDHS. Additionally, PSUs without GPS codes were predominantly concentrated in the Somali region in 2011 and 2016 (see Appendix 2). This may introduce bias to the representativeness of the estimates in the Afar and Somali regions.

In addition, I obtained a shapefile of Ethiopia's administrative boundaries from the United Nations Office for the Coordination of Humanitarian Affairs (OCHA) Ethiopia. This shapefile includes the boundaries of the 11 regional states, 90 zones, and 981 districts (UNOCHA, 2020).

Table 3.2. Number of clusters and women by survey

	EDHS 2000	EDHS 2005	EDHS 2011	EDHS 2016
PSU's Locations				
No. of PSUs with GPS	533	526	569	622
No. of individuals	15,193	13,861	15,730	15,242

Data sources: Ethiopia Demographic and Health surveys 2000, 2005, 2010 and 2016.

3.2.1.2. Study Variables

As mentioned earlier, the EDHS samples are selected using a stratified, two-stage cluster design. To account for the survey design and obtain representative estimates, I applied survey weights as recommended in the DHS manuals. These weights were incorporated into the calculations of the study variables using the srvyr R package.

1) Outcome variable: Total fertility rates (TFR)

The Total Fertility Rate (TFR) is a hypothetical measure used to estimate women's fertility during a specified period. It represents the average number of children a woman would have over her lifetime if she survived all her reproductive years (15-49 years) and experienced the exact age-specific fertility rates observed during that period. TFR can be calculated by adding up all the age-specific fertility rates (ASFR). In the DHS surveys, the individual women's dataset contains information on the birth history of women aged between 15 and 49. To calculate the TFR, I aggregated the number of births and women-years of exposure to pregnancy into seven five-year age groups (15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49 years). This also decreases sampling variability associated with small numbers of annual births due to women in single age groups. Then, $ASFR_a$ can be written as follows;

$$ASFR_{a,c} = \left(B_{a,c}/E_{a,c}\right) \times 1{,}000$$

$a = 1$ - 7 refers to the seven age groups at the time of delivery and c denotes PSUs. $B_{a,c}$ is the number of births to women in age group a in PSU c during a referenced period and $E_{a,c}$ denotes the number of person-years of exposure in age group a in PSU c during a referenced period. In the DHS, TFR is calculated for a reference period of three years prior to the survey and therefore a women can contribute to more than one age category if she moves between two age bands within those three years. To account for this in the calculation women are allowed to contribute to at most two five-year age groups. This was calculated as follows (Croft, 2018):

i) Higher age group: The number of months between the end of exposure (the date of the respondent's interview) and the age group's lower limit is used to calculate the total number of person-years of exposure. When the age group's number of months is greater than or equal to 36 months, the total number of person-years of exposure in the higher age group is 3 years (36 months), while it is zero in the lower age group.

ii) Lower age group: The total number of person-years of exposure in the lower age group is considered when the number of months in the age group for a woman is less than 3 years (36 months), and it is defined as the difference between 36 months and the number of months in the higher age group.

Table 3.3 provides a summary of how to calculate women-year of exposure. Then, TFR for each PSU can be calculated as follows:

$$TFR_c = 5 \times \sum\nolimits_{a,c} ASFR_a$$

Where, $a = 1$ - 7 and refers to the seven five-year age groups and c are the primary sampling units (PSU).

Table 3.3. Example of calculation of women-year of exposure in DHS

Women information	○ Interviewed in **December 2016** ○ Born in **May 1986**		○ Interviewed in **December 2016** ○ Born in **June 1983**	
CMC date of interview	12(Month) X (2016 - 1900)+12 (December) - 1 = **1403**		12(Month) X (2016 - 1900) +12(December) -1= **1403**	
CMC date of birth	12(Month) X (1986 - 1900) + 5(May) = **1037**		12(Month) X (1983 - 1900) + 6(June) = **1002**	
Age in months at the date of interview	1403 - 1037 = **366**		1403 - 1002 = **401**	
Age group at the date of interview	a)	366/60* = **6.1** **60=12 months X 5 years interval*	a)	401/60* = **6.68** **60=12 months X 5 years interval*
	b)	Age group: **30-34 years*** ** 6.1 X 5 years interval=30.5*	b)	Age group: **30-34 years*** **6.68 X 5 years interval=33.4*

Women-years of exposure	**Higher age group:** 30-34	366 – (6 X 60)* + 1 = **7 months** **6th age group X 12months X 5 years intervals* **∴ 7÷12 = 0.58 years**	**Higher age group:** 30-34	401 – (6 X 60)* + 1 = **42 months.** **6th age group X 12months X 5 years intervals* Since the number of months is greater than 36 months, she contributed **36 months** of exposure to age group 30-34 during the period (36 months) **∴ 36÷12 = 3 years**
	Lower age group: 25-29	Since this is less than the total number of months during the period (36 months); 36 - 7 = **29 months** **∴ 29÷12 = 2.42 years**	**Lower age group:** 25-29	**No exposure** to the lower age group during 36 months

2) Explanatory variables

Previous studies have identified various factors that contribute to geographical variations in fertility, including cultural and socioeconomic differences. In line with the objectives of this study, which aim to describe the spatial patterns of selected key proximate and distal determinants of fertility and understand their heterogeneous influences on district-level fertility in Ethiopia, I have built upon this existing research by focusing on two proximate determinants and three distal determinants that have been widely recognized as significant contributors to geographical variations in fertility.

a) Proximate determinants

In 1978, Bongaarts proposed four proximate determinants to explain differences between populations in fertility levels (Bongaarts, 1978). However, I restricted the analysis to include the two main drivers of fertility from Bongaarts proximate determinants, contraception and marriage exposure. I have excluded the other two determinants, namely abortion and postpartum infecundability. This decision is based on previous research conducted in Ethiopia, which has indicated that the differences in the inhibiting effects of contraception and marriage on fertility between regional states in 2011 and 2016 were more substantial compared to the effects of the other two determinants (Laelago et al., 2019).

i) Median age at first marriage

Median age at first marriage is an indicator that uses cumulated single-year percent distributions of age at first marriage. In Ethiopia, childbearing outside of marriage traditionally has not been tolerated (Lindstrom et al., 2009) and therefore, the country has one of the lowest premarital fertility rates in SSA , which never rose above 5% over the past three decades (Clark et al., 2017, Smith-Greenaway and Clark, 2018). Delays in age at first marriage are often partly responsible for the observed fertility declines in many SSA countries (Sobotka, 2017, Ouadah-Bedidi et al., 2012, Bongaarts, 1982, Shapiro and Gebreselassie, 2014, Hertrich, 2017). Although, the prevalence of girls married before the age of 18 in Ethiopia reduced from 59% of females aged 20–24 years in 2005 to 40.3% in 2015 (Gavrilovic et al., 2020, Mekonnen et al., 2018), the prevalence of early marriage and median age at first marriage vary across Ethiopia (Alem et al., 2020). According to the 2016 DHS, the median age at first marriage was the lowest at 15.7 years in Amhara region and the highest at 22.9 years in Addis Ababa (ICF, 2016). The DHS provides the variable of 'age at first marriage(v511)'. Therefore, I calculated the median age at first marriage for each PSU for 2000, 2005, 2011 and 2016. Median values were calculated from the cumulative single-year distribution of age at first marriage at each PSU. To calculate the cumulative single-year distribution, the number of women who were legally or formally married in each single-year age category was used as the numerator. The number of women of all marital statuses, including never-married women, was used as the denominator. Then the median was determined by linearly interpolating between the single-year percentage distributions that correspond to the ages just before and after 50 percent of women have entered into their first marriage (Croft et al., 2018).

ii) Modern Contraceptive Prevalence (mCP) among married women of reproductive age

In the past few decades, increases in the prevalence of contraceptive practice have played a major part in reducing fertility in SSA countries (Cleland et al., 2006, Casterline, 2017). In this study, I used the mCP among married women aged between 15 and 49. According to the Guide to DHS statistics, mCP is defined as the percentage of married women using modern contraception methods (Croft et al., 2018). Among SSA countries, Ethiopia particularly achieved significant progress in enabling women to freely choose the number and timing of their births as evidenced by an increase in the mCP from 6.3% in 2000 to 35.3% in 2016 (ICF, 2016). However, previous studies observed substantial variations in mCPs between regional states in Ethiopia, as an example of 46.9% in Amhara and 11.6% in Afar regions in 2016 (Lakew et al., 2013, Tegegne et al., 2020, Hogan and Biratu, 2004, ICF, 2016). The DHS variable 'use of contraceptive method (V313)' has four categories; no method, folkloric method, traditional method, modern method. Therefore, I calculated mCPs using binary outcomes (1= modern method; 0=other than modern method) for each PSU during 2000, 2005, 2011 and 2016.

b) Distal determinants

Although there remains some debate about which dimension of distal determinants of fertility is most important, urbanisation and improvements in female education are the two most thoroughly studied socioeconomic determinants of fertility (Kebede et al., 2019, Murtin, 2013, May and Rotenberg, 2020, Bongaarts, 2020, Bongaarts, 2017, Robinson, 1963, Gries and Grundmann, 2018). Moreover, according to most DHS reports, Ethiopia has the largest fertility differentials between residential areas (rural and urban areas) and female educational levels among all countries surveyed by the DHS programme (see Appendix 2). In addition, previous studies also showed that geographical areas with common culture and language tend to experience fertility decline at similar time, regardless of the level of socioeconomic status (Watkins, 1987, Watkins, 1990, Bongaarts and Watkins, 1996). Furthermore, the analysis includes the three distal determinants of fertility: a) the proportion of women living in urban areas, b) the proportion of women with secondary or higher education, and c) ethnolinguistic diversity at the zonal level.

i) Proportion of women living in urban areas

Previous studies have shown that differences between urban and rural fertility levels are particularly large in SSA due to the slow pace of fertility decline in rural areas compared with the pace of the decline in rural areas elsewhere in the world (Shapiro and Tenikue, 2017, Garenne and Joseph, 2002, Lerch, 2019). The percentage of people living in urban areas has been extensively used to measure the association between geographical fertility variations and level of urbanisation (Bongaarts, 2020, Bongaarts, 2017, Haque et al., 2019, Singh et al., 2017, Wang and Chi, 2017). Ethiopia is often described as being characterised by substantial differences in fertility levels between urban and rural areas, as evidenced by TFRs of 5.2 and 2.2 in rural and urban areas, respectively, in 2016 (ICF, 2016) Therefore, I calculated proportions of women living in urban areas using binary outcomes (1= living in urban area; 0=living in rural area) for each PSU.

ii) Proportion of women with secondary and higher education

A previous study in low-income countries (LICs) showed that the association between a few years of schooling and the level of total fertility rate (TFR) was inconsistent; however, for women with secondary or higher education, the association was significantly negative (Jejeebhoy, 1995). In SSA countries, fertility differentials also tend to widen as education increases through the secondary level and beyond, suggesting that increased educational attainment at the secondary level and above may hasten the speed of fertility decline in a region (Shapiro, 2012). During the 2000s, Ethiopia experienced a relatively large drop in the fertility of women with secondary or higher education, as both parity-specific limitation and postponement behaviours among those women increased (Towriss and Timæus, 2018). The 2000 and 2005 EDHS defined 'secondary and higher education' as seven years of schooling or more, while the 2011 and 2016 EDHS defined it as nine years of schooling or more. In this thesis, I define 'secondary and higher education' as nine years of schooling or more using the DHS variable 'Highest year of education (V107)'. Then, I calculated proportions of women with secondary and higher education using binary outcomes (1 = woman with secondary or higher education; 0 = women with no schooling or primary education) for each PSU during 2000, 2005, 2011, and 2016.

iii) Ethnolinguistic homogeneity at Zonal areas

The diffusionist approach argues that cultural heterogeneity, resulting from different ethnicities and languages, can hinder the equal diffusion of attitudes and information that support modern reproductive ideas and behaviours (Cleland and Wilson, 1987, Watkins, 1987). To assess the similarity or diversity in ethno-languages between neighbouring districts in Ethiopia, I utilized the standardized index of diversity at the zonal level. The UN OCHA shapefile consists of 90 zones, with an average of 10 woredas per zone. The index of diversity, also known as the 'entropy index,' has previously been employed to measure the impact of ethnic or cultural diversity on fertility behaviour (Hogan and Biratu, 2004) . The entropy index is defined as (White, 1986);

$$ID_z = -\sum_{g=1}^{g=G} dv_{zg} \ln (dv_{zg})$$

Where z = 1 to 90 (The number of zonal areas for this study), and g refers to the number of ethnolinguistic groups. Therefore,

$$dv_{zg} = \frac{N_{zg}}{N_z}$$

where N_{zg} is the number of persons in the g^{th} ethnic group in the z^{th} local area, N_z denotes the total population size of the z^{th} local area, and G refers to the total number of ethnic groups in z^{th} local area. While lower values of the index (close to zero) indicate similarity in ethnolinguistic composition at zonal areas, larger values of the index show ethnolinguistic diversity at zonal areas. The EDHS variable 'Ethnicity (v131)' contains approximately 90 ethnic groups in Ethiopia for individuals. Therefore, I calculated the index of ethnolinguistic diversity at the zonal level, resulting in PSUs within the same zone sharing the same value for the index, during 2000, 2005, 2011, and 2016.

3.2.1.3. Data Analysis

The Demographic and Health Survey (DHS) Program has provided georeferenced data on important demographic and health indicators in sub-Saharan Africa over the past twenty years, and this data has been widely used in recent years. The DHS Spatial Interpolation Working Group, which assessed properties of various Spatial Interpolation (SI) methods, suggested the Bayesian model-based geostatistical (MBG) technique as the most suitable for producing interpolated surfaces (Burgert-Brucker et al., 2016b, Gething et al., 2015). The DHS Spatial Analysis Reports (SAR) 17 and 19 recently produced estimates of health outcomes at the second subnational administrative level (Admin 2) using a stochastic partial differential equation (SPDE) with the integrated nested Laplace approximation method (INLA), because health programs are frequently implemented at the Admin 2 level (Mayala et al., 2019a, Fish et al., 2020). In this thesis, I utilise a similar Spatial Interpolation (SI) method, namely MBG implemented through the INLA-SPDE approach. However, I generate estimates at the third subnational administrative level (Admin 3) in Ethiopia. This is because the District-level Decentralization Programme (The DLDP) in 2001/02 decentralized health planning and service delivery to the third subnational administrative level (Admin 3) in Ethiopia (Garcia and Rajkumar, 2008). In this section, I provide an overview of the MBG model and outline the ten steps I used to predict Total Fertility Rate (TFR) and key selected proximate and distal determinants for 981 districts in Ethiopia using the four EDHS datasets (2000, 2005, 2011 and 2016).

It is important to note that the methods employed in this thesis for estimating the values of study variables in small and unsampled areas (districts) differ from the conventional approach of small area estimation (SAE). In SAE, direct survey data is collected from a sample of individuals within each small area, and these estimates are combined with auxiliary information from a larger area, such as the entire nation, to enhance the precision of the small area estimates (Alho, 2001). For instance, this approach was adopted to estimate total fertility rates for over 5,000 municipalities in Brazil by using the 2000 Brazilian Census data (Schmertmann et al., 2013). They used a sophisticated small area estimation method by applying empirical Bayes methods, but this method has a limitation in estimating fertility levels for unsampled areas. Moreover, it should be noted that although Ethiopia's population census data provide fertility rates at the district level, the two most recent population censuses in Ethiopia were conducted in 1998 and 2007. Therefore, analysing the change in recent geographical patterns of fertility using Ethiopia's census data with the conventional approach is challenging. Instead, alternative types of data, including more frequently collected household sample surveys, can be used to estimate population indicators. In contrast to the conventional approach, the methods used in this thesis take alternative approaches, utilising model-based techniques to estimate the study variables in small and unsampled areas. Rather than relying solely on direct survey data, these methods leverage additional information, such as spatial or temporal patterns, to generate subnational-specific estimates (Mercer et al., 2015). By incorporating a wider range of data and employing advanced statistical modelling techniques, these model-based approaches have the potential to improve the precision and reliability of estimating the values of variables in small areas for both sampled and unsampled locations (Aheto and Dagne, 2021).

1) Overview of Model-based geostatistics using INLA-SPDE model approach

Geostatistics is a field of study that focuses on the analysis of continuous processes in space. Its primary objective is to predict values for unobserved areas, a process often referred to as "spatial interpolation." A common example of spatial interpolation is the estimation of temperature distribution across an area. Since observations of the variable are only available at a limited number of locations, statistical methods are used to estimate the variable for the entire study area. Model-based geostatistics (MBG) is an approach that applies general statistical modelling principles to make probabilistic inferences about spatially continuous phenomena using data collected at a limited number of georeferenced locations. The MBG model is a type of generalized linear mixed model (GLMM) that enhances the flexibility of generalized linear regression models by incorporating spatial variability through a multivariate normal distribution (Diggle et al., 1998). The MBG model encompasses three categories of parameters: fixed effects, random effects, and a simple Gaussian noise term, which are similar to those employed in standard nonspatial linear models (Andres et al., 2018, Mayala et al., 2019b, Burgert-Brucker et al., 2016a). The MBG model can be specified as:

$$\underbrace{g\big(\mathrm{E(y)}\big)}_{\substack{\text{Expected}\\ \text{outcome}}} = \underbrace{X\beta}_{\substack{\text{Fixed}\\ \text{Effects}}} + \underbrace{u(s)}_{\substack{\text{Random}\\ \text{Effects}}} + \underbrace{\varepsilon}_{\text{Noise}}$$

Where, $g(\cdot)$ is the link function that links the outcome variable to the expected value, $\mathrm{E(y)}$, to the linear predictors, $X\beta,\ u(s)$ and ε. β refers to the matrix of coefficient and X denotes the matrix of covariates (fixed effects). $u(s)$ is a geostatistical random effect (random effects) and ε is the uncorrelated residual error (Noise). These three components are introduced as follows.

a) Fixed effects

In the context of the MBG model, unsampled variations can be accounted for by incorporating fixed effects. It is crucial to carefully select the appropriate set of fixed effects or covariates in order to improve the estimation accuracy of the model. However, it is important to note that there is a distinction between producing the "best possible map" and producing a "standardized map" for a specific country, as highlighted by the DHS Spatial Interpolation Working Group (Burgert-Brucker et al., 2016b). While producing the best possible map can enhance map accuracy, it may hinder direct comparisons between countries or over time if the availability of covariate datasets is inconsistent. Therefore, to adopt a comparative approach, it is advisable to use covariates that are consistently available across countries or over time. The DHS survey collects various variables from survey clusters that could potentially be used as fixed-effect covariates in the model. However, these variables are not available throughout the entire study area, rendering them unsuitable for spatial interpolation (Burgert, 2014). As an alternative, previous studies have utilized geospatial covariates obtained from publicly available remote sensing sources, such as land surface, temperature, and average monthly rainfall, for mapping disease prevalence or environmental factors (Giorgi et al., 2021, Huang et al., 2017, Reiner Jr et al., 2020). While these covariates can help explain variations in data for certain communicable diseases like malaria, they could only account for a small portion of the variation in socially determined health outcomes, such as HIV and fertility (Mayala et al., 2020).

b) Random effects

Spatial variation that cannot be explained by covariates can be accounted for through the random effects component. Geostatistical data are measurements about a spatially continuous phenomenon that is collected at a finite set of sampled locations. Suppose $Z(s_1), \ldots, Z(s_n)$ are observation of a variable Z at sampled locations $s_1, \ldots, s_n$. Therefore, geostatistical data are often assumed to be a partial realisation of a random process.

$$[Z(s): s \in D \subset \mathbb{R}^2]$$

where, D is a fixed subset of $\mathbb{R}^2$ and the spatial index s varies continuously over D. $Z(s)$ is observation of the variable Z at space s. As $Z(s)$ is often observed at a finite number of locations, we need to make inferences about the characteristics of the spatial process underlying the observed data, such as mean and variability of the process. These characteristics are useful for the prediction of the process at unobserved locations (Moraga, 2019).

In geostatistics, it is commonly assumed that this random process follows a Gaussian distribution, resulting in a spatial Gaussian Process (GP) (Cressie, 2015). In addition, there are two main characteristics often assumed for the GP. The first characteristic is stationarity, which implies that the covariance between two observations remains constant even when their locations are spatially shifted. The second characteristic is isotropy, which means that the spatial correlation between two observations is solely determined by the distance between them and not by their specific directions. These two characteristics simplify the modelling of the geostatistical process, as the spatial correlation can be modelled using an appropriate covariance function (Salvati et al., 2020). Furthermore, a Gaussian random field (GRF) is a collection of random variables in a continuous domain, where any finite set of random variables has a multivariate normal distribution with a mean of zero and a joint covariance structure. It is important to note that a one-dimensional GRF is equivalent to a GP. A commonly used and very flexible covariance function in geostatistics is the Matérn covariance function, which is described as (Cressie, 2015)

$$\mathrm{C}ov(Z(\boldsymbol{s}_i), Z(\boldsymbol{s}_j)) = \frac{\sigma^2}{\mathbf{2}^{v-1}\Gamma(v)}\left(k\|\,\boldsymbol{s}_i - \boldsymbol{s}_j\|\right)^v K_v\left(k\|\,\boldsymbol{s}_i - \boldsymbol{s}_j\|\right)$$

Here, $\Gamma(\cdot)$ is the gamma function and $\|\ \boldsymbol{s}_i - \boldsymbol{s}_j\ \| \in \mathbb{R}$ denotes the distance between locations $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$, σ^2 denotes the spatial variance. $\kappa > 0$ is a scaling parameter related to the spatial range $r = \frac{\sqrt{8v}}{k}$ that is the distance at which the spatial correlation becomes almost null and v is a smoothness parameter. $K_v(\cdot)$ is the modified Bessel function of the second kind. It is normally difficult to learn about the smoothness parameter v, and therefore it often fixes this parameter. This study follows this convention by setting v =1 (Wilson and Wakefield, 2020). The Matérn covariance function was used in the recent DHS spatial analysis reports to estimate health indicators at the desired administrative level (Burgert-Brucker et al., 2016b, Mayala et al., 2019b). Figure 3.1 illustrates the Matern covariance function.

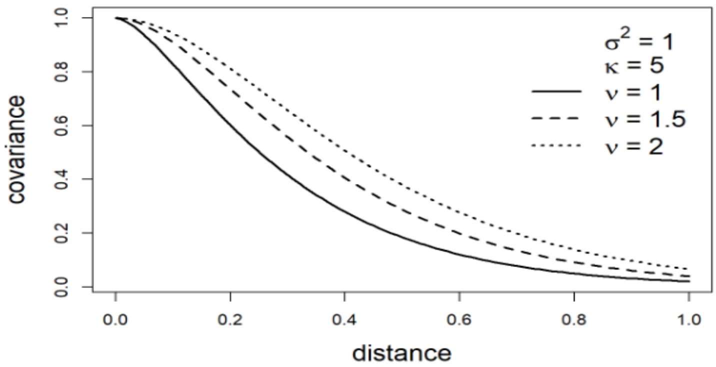


Figure 3.1. Covariance functions of Matern models
Source: Geospatial Health data: Modeling and Visualization with R-INLA and Shiny (Moraga, 2019)

To assess spatial correlation of study variables of DHS data, spatial correlation was modelled by a variogram, which is a commonly used tool in geostatistics to visualise the spatial structure (Diggle and Ribeiro, 2007). The variogram (semi-variogram) describes the extent to which nearby locations exhibit similar values by measuring the semi-variance. In standard statistics, correlation can be estimated from a scatterplot when multiple data pairs are available. However, in the case of spatial correlation between two observations of a variable $Z(s)$ at locations s and $s + h$, where h represents the separating distance (lag), estimation is not possible due to the presence of only a single pair of observations. Geostatistics overcomes this limitation by adopting the assumption of stationarity, which enables us to assume that the covariance between observations only depends on the lag h, regardless of their location. Then the covariance between two observations can be written as $Cov(Z(s),\ Z(s + h))$. By the assumption of stationarity, a variogram $\gamma(h)$ can be defined as:

$$\hat{\gamma}(h) = \frac{1}{2n}\sum_{i=1}^{N}\{Z(s_i + h) - Z(s_i)\}^2$$

Where N represents the number of pair points of PSUs separated by lag distance h used to estimate the value of $\hat{\gamma}(h)$. This collection of points is commonly referred to as the variogram cloud, illustrating the semivariances between all pairs of points. The purpose of this function is to measure the similarity in attributes between neighbouring observations at a lag distance h. In the semi-variogram, this implies that as the spatial separation between observations grows, the semi-variance is expected to increase since nearby observations tend to share more similarities compared to those that are far apart (Figure 3.2). Some important characteristics of the variogram are as follows:

1) Range refers to the critical distance beyond which there is no longer any spatial correlation observed in the variogram.
2) Sill represents the maximum value of the semi-variance, indicating the extent of variability in the absence of spatial correlation.
3) Nugget represents the semi-variance as the separation distance approaches zero. It quantifies the variability at a point that cannot be explained by spatial structure.

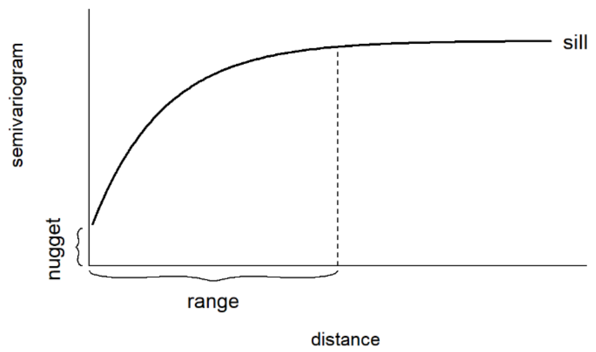


Figure 3.2. Typical diagram of the geographical semi-variogram
Source: Geospatial Health data: Modeling and Visualization with R-INLA and Shiny

The first stage of a geostatistical analysis is to check the evidence of spatial correlation. I first compared the empirical semi-variograms to a 95% pointwise envelope based on 1,000 Monte Carlo simulations (Diggle and Ribeiro, 2007). If the empirical semi-variograms lies outside the Monte Carlo envelope, there is evidence of spatial correlation.

c) Noise

The remaining variations that are not captured by the fixed and random components are represented by a simple Gaussian noise term, which is commonly applied in non-spatial linear models.

In this study, the Integrated Nested Laplace approximation (INLA) technique from the R-INLA package was used to apply the Model-Based Geostatistics (MBG) model using a Stochastic Partial Differential Equations (SPDE) approach (Rue et al., 2009). Integrated Nested Laplace approximation (INLA) is a computationally efficient method for performing approximate Bayesian inference in latent Gaussian models (Rue et al., 2009). Latent Gaussian models are family of wide and flexible class of models ranging from generalized linear (mixed) models (GLMMs) to spatial and spatio-temporal models. Spatial and spatio-temporal models can also be estimated, since the spatial model is essentially a form of GLMM with random effects of spatial structure. The INLA approach offers fast and reliable calculations of the posterior marginal distribution compared to the Markov Chain Monte Carlo (MCMC) algorithm, which involves dense matrices that increase computation time (Mayala et al., 2019b). While INLA was originally developed for discrete space, recent developments have connected INLA with the Stochastic Partial Differential Equation (SPDE) framework, enabling its use in continuous space analysis. This approach involves discretizing the continuous space into a large number of discrete spaces using a constrained refined Delaunay triangulation, commonly known as a mesh, over the area of interest. I utilized the INLA-SPDE approach for estimating study variables at 981 local areas in Ethiopia (Lindgren and Rue, 2015, Blangiardo et al., 2013). The INLA-SPDE approach was implemented using the R package R-INLA (http://www.r-inla.org/).

2) Ten Steps to use a Bayesian model-based geostatistics (INLA-SPDE) to small areas estimation of study variables with the EDHS data (Huang et al., 2017)

In this thesis, I assumed that the spatial process for the selected study variables follows a continuous process represented by a Gaussian Random Field (GRF). To model and predict these variables at unobserved locations, I used the Stochastic Partial Differential Equations (SPDE) framework implemented in the R-INLA package (Lindgren et al., 2011, Moraga, 2019, Blangiardo and Cameletti, 2015). The SPDE approach, proposed by Lindgren et al. (2011), provides an approximate solution to the SPDE by employing the Finite Element method (Lindgren et al., 2011). This method involves creating a triangulated mesh that divides the spatial domain (D) into non-intersecting triangles. Defining the SPDE model in the R-INLA package requires several steps, with one crucial step being the creation of a mesh over the study region to calculate an approximation to the solution, which represents the spatial process. Here, I summarise steps required to fit the SPDE model in the R-INLA package. Krainski et al. provide a detailed description of the SPDE model (Krainski et al., 2018).

- Step 1: calculating the mean value of study variables at PSUs

First, mean values of the study variables were calculated at each Primary Sampling Unit (PSU), taking into account survey-specific weights as outlined in the DHS manuals provided by EDHS. It is important to note that within PSUs, the sample sizes might be small, potentially leading to extreme values. However, the geostatistical approach can mitigate this issue by borrowing information from neighbouring areas, resulting in the smoothing or shrinking of extreme values and capturing the level of uncertainty (Gething and Burgert-Brucker, 2017, Burgert-Brucker et al., 2018, Burgert-Brucker et al., 2016b).

- Step 2: Building the mesh

The second step involves creating a neighborhood structure, referred to as a mesh, for the continuous data. This can be achieved by implementing a constrained Delaunay triangulation, where the starting vertices are based on the observation locations (Moraga, 2019). In practice, constructing the mesh involves finding a balance between approximation accuracy and computational costs.

First, I obtained a triangulation using the initial vertices located at the PSU locations corresponding to the years 2000, 2005, 2011, and 2016, specifically vertices 533, 526, 569, and 622. To ensure the quality of the triangulation, additional vertices were added based on the (1) cutoff, (2) offset, and (3) max.edge parameters of the inla.mesh.2d() function in the R-INLA package (Blangiardo and Cameletti 2015; Lindgren, Rue, and Lindstrom 2011). In the the R-INLA package, The cutoff parameter is used to prevent the creation of excessively small triangles around the PSU locations, while the offset parameter determines the size of the inner and outer extensions surrounding the PSU locations. The max.edge parameter sets the maximum length for triangle edges in the inner domain and outer extension. In order to maintain a finer approximation at a reasonable computation time and cost, I limited the number of vertices to approximately 4,000 (Figure 3.2).

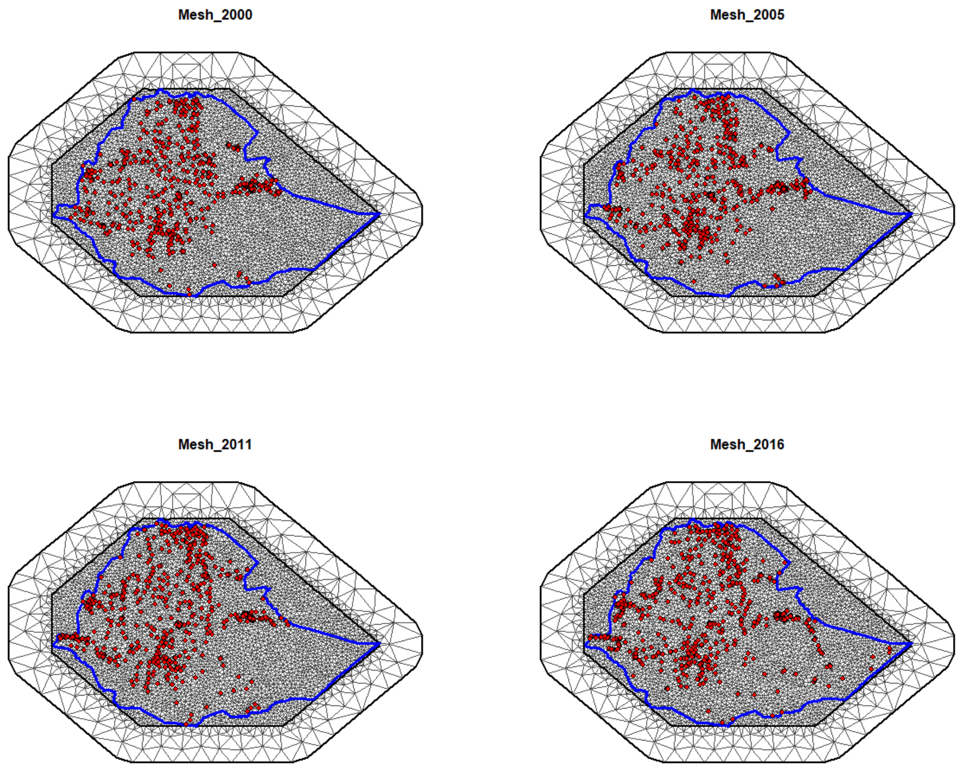


Figure 3.3. The Ethiopia triangulation with 3952 vertices for 2000, 3924 vertices for 2005, 3946 vertices for 2011 and 3930 vertices for 2016.
Note: The red dots mark location of Primary sampling units (PSUs)

- Step 3: Defining the SPDE model in the mesh

Here the Matérn covariance function was used.

- Step 4: Prior

The choice of prior in a geostatistical model can be either informative or non-informative. In this study, I opted for the default setting of non-informative priors (Poggio et al., 2016). By default in R-INLA package, the intercept has a Gaussian prior with mean and precision equal to zero. Coefficients of the fixed effects also have a Gaussian prior by default with zero mean and precision equal to 0.001. The prior on the precision of the error term is a Gamma distribution with parameters 1 and 0.00005 (Blangiardo and Cameletti, 2015). I fit the model by calling inla() and using the default priors in R-INLA.

- Step 5: Fit the hierarchical model

I defined the likelihood family for the model response variables based on their nature and characteristics. The variables TFR (Total Fertility Rate), median age at first marriage, and ethnolinguistic diversity index were treated as continuous variables and assumed to follow a Gaussian (normal) distribution. On the other hand, the variables mCP (modern contraceptive prevalence), proportion of women living in urban areas, and proportion of women with secondary or higher education were treated as binomial variables and assumed to follow a binomial distribution. The binomial distribution is appropriate for variables that represent a count or proportion with a fixed number of trials (in this case, the total number of women surveyed). These choices of likelihood distributions allow for appropriate modelling of the variability and distributional characteristics of the variables under consideration. Table 3.4 provides an overview of the likelihood families used for each variable.

Table 3.4. Study variables and probability distributions

Study variables	Type
▪ Total Fertility Rate	Continuous variable
▪ Median age at first marriage	
▪ Ethnolinguistic diversity (Entropy index)	
▪ Modern contraceptive prevalence (mCP) (1= using modern contraceptive method, 0= not using)	Binary variable

▪ % women with secondary or higher education (1= Secondary or higher, 0 = lower than secondary)	
▪ % Urban population (1= living in urban area, 0= living in rural area)	

- Step 6: Construct a hierarchical model

I employ a geostatistical model to estimate values of study variables in unsampled areas of Ethiopia under the assumption that fertility and key selected determinants occur continuously in space. In the case of fertility, fertility at spatial location i, follows a zero-mean Gaussian process with a Matérn covariance function, and the mean is determined by the sum of an intercept and a spatially structured random effect (Table 3.5). For continuous variables, such as TFR and median age of first marriage, I used a Gaussian spatial model to estimate the values at each spatial location i, after having checked normality in the total fertility rate, the median age at first marriage and the ethnolinguistic diversity index (Appendix 5). The model assumes a Gaussian distribution for the response variable, with a mean determined by spatial covariates and a spatially structured random effect. The variance of the response variable is also accounted for in the model. For binary variables, I used a binomial spatial model with a logit link function. This model allows for the estimation of the probability of the event (e.g., proportion of women with secondary education) at each spatial location i, taking into account the spatial structure and covariates. The total number of women sampled at each spatial location, denoted as N_i, is incorporated in the model to account for the sample size variation (as shown in Table 3.5).

To improve the predictive accuracy of the model, it is important to select the most appropriate set of covariates. In the case of this study, the variables collected from DHS surveys at each cluster are not suitable for spatial interpolation because they are not observed throughout the entire mapping region (Burgert, 2014). Therefore, alternative geospatial covariates obtained from publicly available remote sensing sources can be utilized. These covariates, such as land surface temperature, enhanced vegetation index, and average monthly rainfall, have been commonly used in studies related to disease prevalence or environmental mapping (Giorgi et al., 2021, Huang et al., 2017, Reiner Jr et al., 2020). While these covariates may be associated with environmental variables and diseases like soil carbon and malaria, their association with fertility and key determinants of fertility is less likely. Furthermore, it is important to clarify that the main objective of this study is to explore the spatial variations in fertility rather than investigating the specific changes in the estimated study variables at the district level associated with a one-unit increase in covariates. In light of this, the intercept term (β_0) is treated as a fixed effect in the model, representing its overall impact on fertility across the study area. On the other hand, the spatial random effect ($u(S_i)$) is included in the model to capture the unexplained spatial variation in fertility (Table 3.5). By considering β_0 as a fixed effect and incorporating $u(S_i)$ as a spatial random effect, this approach helps remove variation in the estimates that could potentially influence the exploration of associations between the outcome variable (fertility) and the explanatory variables for the study objectives (3, 4, and 5). This allows for a more focused analysis on understanding the spatial patterns and variations in fertility, without being confounded by the specific effects of the covariates. Additionally, it is worth noting that this study employs a cross-sectional analysis, meaning that a separate model is fitted for each year (2000, 2005, 2011, and 2016).

Table 3.5. Bayesian hierarchical models for continuous and binomial variables

Continuous variable	Binary variable
$Y_i \sim Normal(\mu_i, \sigma^2)$, $i= 1, 2, \ldots, n$	$Y_i \sim Binomial(N_i, \pi_i)$, $i= 1, 2, \ldots, n$
$\mu_i = \beta_0 + u(S_i) + \varepsilon_i$	$logit(p_{i)} = \beta_0 + u(S_i) + \varepsilon_i$
$u(S_i) \sim GP(0, \Sigma)$	$u(S_i) \sim GP(0, \Sigma)$

where;

- Y_i denotes either continuous or binomial study variables for spatial location t i
- μ_i is the mean, representing the underlying mean for spatial location i
- π_i is the probability, representing the underlying prevalence for spatial location i
- β_0 denotes the intercept
- $u(S_i)$ = is the spatial error accounting for spatial autocorrelation between data points. It is modeled as Gaussian Process with a zero mean and the spatially structured covariance matrix Σ based on the Matérn covariance function
- ε_i is an unstructured random error term known as nugget effect.

  - Step 7: Estimate the posterior distribution of the parameters

Here, parameters of the matern function () and the marginal distribution of the intercept were estimated using INLA-SPDE.

  - Step 8: Predict the study variables and calculate their posterior marginal distribution

I predicted values of study variables at 145,978 spatial locations, which is approximately equivalent to 3km^2 grid in Ethiopia.

  - Step 9: Spatially modelled map surface

The MBG models produce estimates of the study variables at 145,978 spatial locations, which is approximately equivalent to 3km^2 grid in Ethiopia. The model provides two separate surface maps.

a) Point estimate surface: This map provides the modelled point estimate at 145,978 spatial locations based on geo-referenced data from the DHS. This estimate shows the expected value of the indicators at 145,978 spatial locations.

b) Uncertainty surface: A map of uncertainty shows the level of uncertainty related with the expected values of the point estimate surface by plotting the 95 % confidence intervals (CI) for each spatial location.

- Step 10: Calculate area-average for each district

Aggregation from the point estimate to polygons or districts is one of the most important ways in which modelled surfaces can be manipulated to inform policies and programmes. There are two main methods commonly used for this type of aggregation (Burgert-Brucker et al., 2016b):

a) Simple mean zonal statistics: Values for the polygon are determined by taking the average of all point estimates within the polygon.

b) Population weighted mean statistics: Apply the same method, but consider the population of each grid square and how it affects the estimate for the whole study area.

The simple mean zonal statistics method, as mentioned by Burgert-Brucker et al. (2018), can be used to estimate values at the aggregated polygons (Burgert-Brucker et al., 2018). This method calculates the average of the values within each polygon without considering the population size or distribution. n previous studies, researchers have used population-weighted mean statistics by incorporating gridded population estimates from projects like WorldPop. However, Burgert-Brucker et al. (2016b) emphasized the importance of selecting the appropriate reference population for the denominator estimation, considering factors such as gender and age-groups (Burgert-Brucker et al., 2016b). For this study, the denominator for estimating the aggregated values should be women aged 15-49 between the years 2000 and 2016. While population density and count estimates for individuals are available from the WorldPop project, there is a scarcity of population data specifically for women of childbearing age (15-49 years old) during 2000-2016.

Furthermore, considering the differences in population density between administrative level 3 (Admin 3) areas, it is expected that the variations within each Admin 3 area would be much smaller compared to the variations within Admin 1 and Admin 2 areas. This is illustrated in Figure 4. Based on this understanding, I assumed a constant population density for women aged 15-49 within districts. This assumption is justified by the fact that districts are smaller and more homogeneous areas compared to higher administrative levels. To estimate aggregated values within districts, I used the simple zonal statistic method, which calculates the average value within each district without taking population size into account. This approach is appropriate considering the assumption of constant population density within districts.

Additionally, to account for the complex survey design of the Ethiopian Demographic and Health Surveys (EDHS), I applied survey weights as outlined in the DHS manuals. These weights are used to adjust for the sampling design and ensure that the estimates are representative at the national and subnational levels, taking into account stratification, clustering, and the different probabilities of selection for each sampled individual or primary sampling units (PSUs).

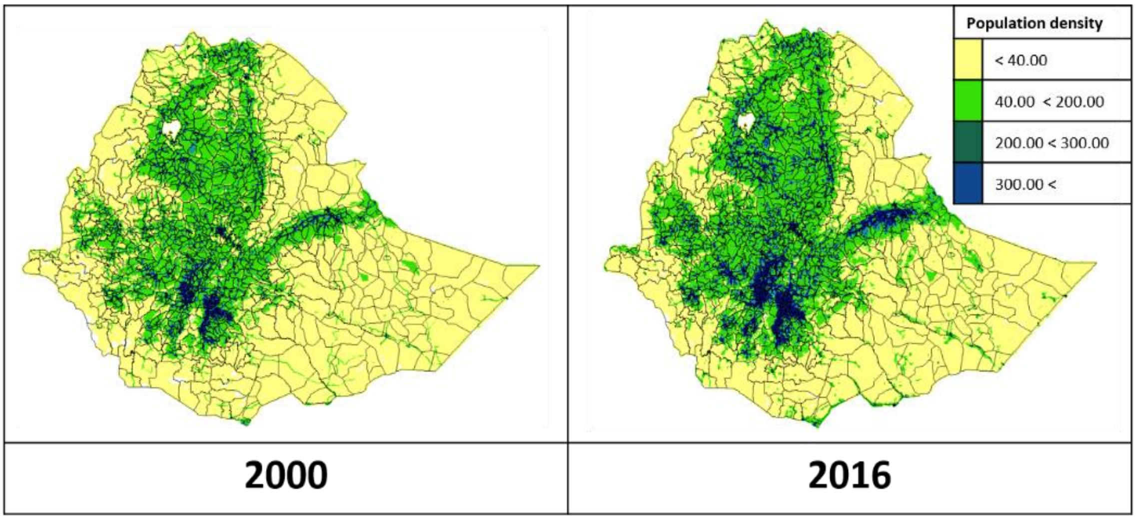


Figure 3.4. Estimated population density estimates (1 km resolution) in 2000 and 2016.
Note: Gray lines refer to district boundary. (Data source: The WorldPop Project)

Hence, for each district $\boldsymbol{A}_j$, the average was computed using simple mean: the area is computed using;

$$\widetilde{\boldsymbol{y}_J} = \frac{\sum_{s_i \in \boldsymbol{A}_j} y(s_i)}{s_i \in \boldsymbol{A}_j}$$

where, $s_i \in \boldsymbol{A}_j$ denotes the total number of predicted locations inside district A_j and $y(s_i)$ denotes predicted values for y variables at s_i location.

Lastly, I compare the geostatistical indicator estimates and directly calculated indicators from the EDHS between 2000 and 2016 to examine the goodness of fit between model-based and observed estimates at the regional level. Model validation was performed only at the regional level since the Ethiopian Demographic and Health Survey (EDHS) was originally sampled to represent the national and regional levels.

3.2.2. Objective 2: To describe and explore spatial and temporal patterns of study variables in 2000, 2005, 2011 and 2016

3.2.2.1. Data and study variables

I used the modelled fertility and key-indicator estimates at 981 districts between 2000 and 2016 that I obtained from the study objective 1.

3.2.2.2. Data Analysis

I visualised fertility and key-indicator estimates at 981 districts between 2000 and 2016 in Ethiopia's to explore spatial and temporal patterns of study variables.

Figure 3.4 illustrates the proposed geostatistical modelled prediction process from the DHS data inputs to the outputs at the local levels.

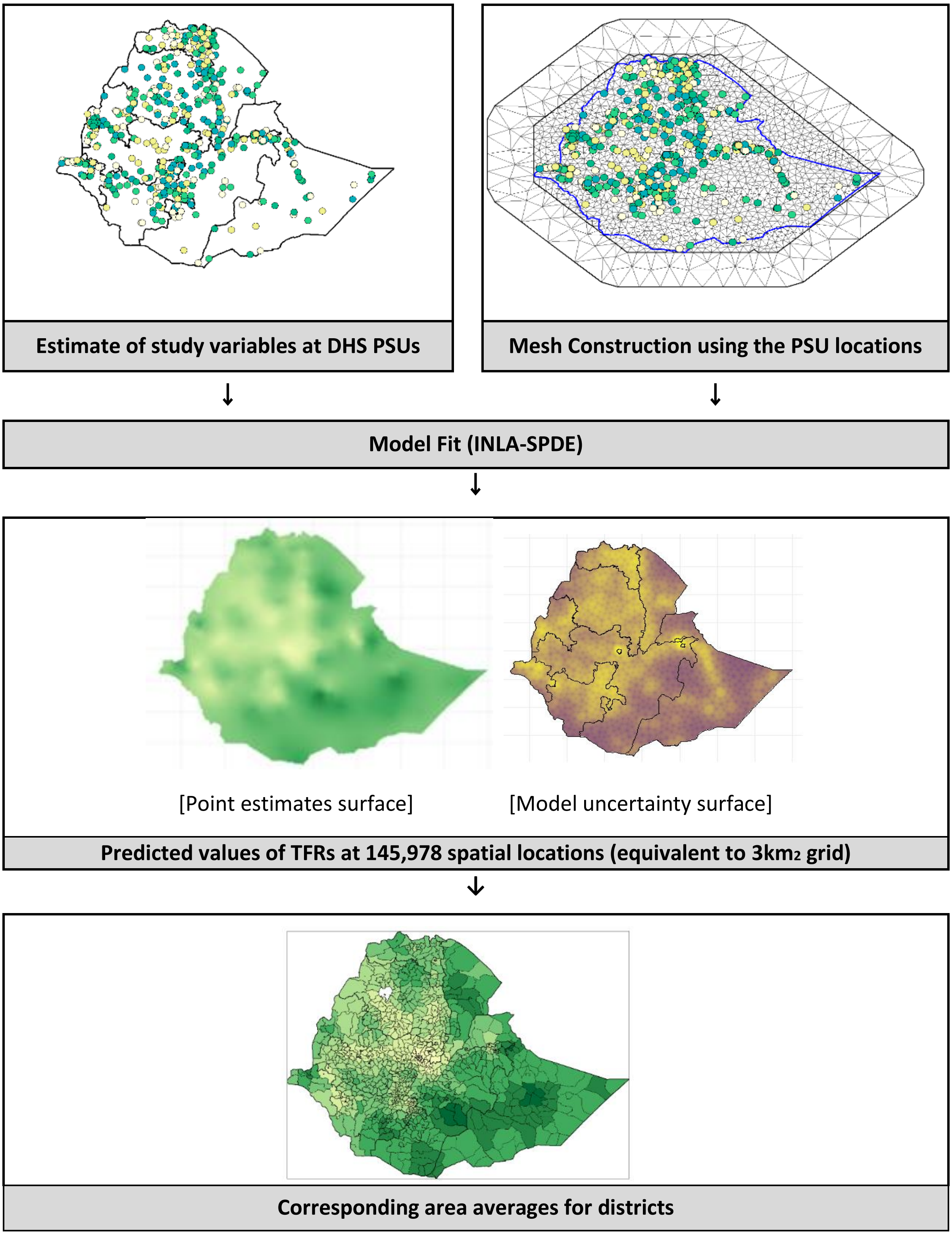


Figure 3.5. Geostatistical modelled prediction process

3.3. Paper 2: What determines geographical variations in fertility at the district level?

In paper 2, I utilise various spatial models to examine the spatial effects of key selected determinants on fertility at the district level in Ethiopia, using the district-level data obtained from paper 1. Firstly, I compare non-spatial and spatial generalized linear models to determine the existence of spatial dependency in fertility in Ethiopia. Next, I investigate the spatial heterogeneity in the relationships between Total Fertility Rates (TFRs) and both proximate and distal determinants in Ethiopia from 2000 to 2016. This research paper addresses study objectives 3, 4, and 5 as outlined in Table 3.1.

3.3.1. Data and study variables

I use the 2000, 2005, 2011, and 2016 Ethiopian Demographic and Health Surveys (EDHS) to assess the spatial autocorrelation of fertility at the primary sampling units (PSUs) across the study periods. Subsequently, I utilise the fertility and key-indicator estimates for the 981 districts between 2000 and 2016, which were obtained from the study objective 1.

3.3.2. Data Analysis

I employed various spatial models to examine the spatial autocorrelation and heterogeneity of fertility at the district level in relation to selected proximate and distal determinants. Since I conducted a cross-sectional analysis, I fitted separate spatial models for each year instead of using a single model with effects for the year.

3.3.2.1. Objective 3: To assess effects of key selected distal and proximate determinants on geographical variations in fertility at the district level between 2000-2016 with the non-spatial linear model

To assess the direct effect of indicators on fertility levels within the same region, I employed a non-spatial regression model. The model is formulated as follows:

$$y(TFR)_j = \beta_0 + \sum_{m=1}^{N} \beta_m X_{mj} + \varepsilon_j$$

where j denotes the j^{th} districts, where j = 1 to 981; y_j is the outcome variable, TFR at district j; β_m is the regression coefficient for explanatory variable . m denotes the selected explanatory variables, including mCP, Median age at first marriage, proportion of urban population and women with secondary and higher education and ethnolinguistic homogeneity at Zonal areas. X represent the value of explanatory variable m at district j. ε_j is an error term for the regression equation. However, this non-spatial regression model does not consider any form of spatial effects between different places.

To examine spatial autocorrelation of TFR at the primary sampling unit (PSU) level, I conducted an analysis of TFR at the PSU level between 2000 and 2016 by using semi-variance analysis. Specifically, I present a) TFR at the PSU level and b) the variogram cloud, and c) the semi-variogram plots for TFR by using an exponential model.

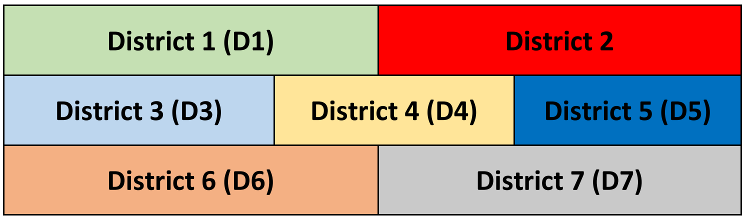
3.3.2.2. Objective 4: To assess spatial dependency of district-level fertility by using the spatial linear model

As explained earlier, according to the diffusion theory, changes in fertility rates spread from one area to another through social interactions and the diffusion of ideas and behaviours. This can cause the tendency of similar fertility rates to cluster together in geographic space regardless of socioeconomic conditions. In the context of fertility, it means that districts in close distance are likely to have similar fertility rates due to the diffusion of fertility behaviours. In spatial statistics, spatial lag regression is often used to account for spatial autocorrelation by incorporating neighbouring values into the regression model. It allows us to examine the relationship between fertility rate in a district and the fertility rates of neighbouring districts, capturing the diffusion process explicitly. A spatial lag regression model includes a lagged dependent variable (fertility rate of neighbouring areas) as an independent variable in addition to other relevant independent variables. This accounts for the spatial autocorrelation of fertility rates and allows us to estimate the effects of neighbouring fertility rates on a given district's fertility. Therefore, if the coefficient of the lagged dependent variable is positive and statistically significant, it suggests that an area's fertility rate is influenced by the fertility rates of its neighbouring areas, supporting the diffusion theory.

Existing fertility studies have utilised the SLM to model the spatial diffusion of fertility, incorporating an autocorrelation coefficient on fertility (Vitali and Billari, 2017, Montgomery and Casterline, 1993, Goldstein and Klüsener, 2014). The SLM is particularly suitable for capturing spatially diffusive processes in the outcome variable, as diffusion processes tend to spread among people over space (Fingleton and Le Gallo, 2008). In order to test the spatial autocorrelation of total fertility rate (TFR) at the district level, I employed a spatial linear model and compared the results with a non-spatial linear model. The SLM is advantageous in accounting for the proximity of neighbouring spatial units. To explore the spatial dependence of TFR, I employed the following SLM:

$$y(TFR)_j = \beta_0 + \rho \sum_{cn=1}^{N} W_j y_{cn} + \sum_{m=1}^{N} \beta_m X_{mj} + \varepsilon_j$$

ρ is the spatial lag term that reflects the strength of spatial autocorrelation in fertility. *cn* refers to the connected neighbouring districts and y_{cn} is the TFR of the neighbouring district. X represent the value of explanatory variable *m* at location j. W_j denotes the spatial weight for a given district j. For spatial weights, I employed a first order queen contiguity approach that assigns a binary spatial weight (0,1) to any connected neighbouring districts (Anselin and Arribas-Bel, 2013, Anselin and Rey, 1991). For example, if a district is connected to three districts, then it will have three links in the weight matrix. (Figure 3.5). I used the contiguity approach, instead of a distance-based approach. This is because the different sizes of districts can result in unequal representation of spatial connectivity.



$$\begin{array}{c c} & \begin{array}{ccccccc} \mathbf{D1} & \mathbf{D2} & \mathbf{D3} & \mathbf{D4} & \mathbf{D5} & \mathbf{D6} & \mathbf{D7} \end{array} \\ \begin{array}{c} \mathbf{D1} \\ \mathbf{D2} \\ \mathbf{D3} \\ \mathbf{D4} \\ \mathbf{D5} \\ \mathbf{D6} \\ \mathbf{D7} \end{array} & \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

Figure 3.6. Diagram of the first queen contiguity weight matrix

Although some reports use spatial error model (SEM) to test spatial autocorrelation, SEM assumes that the spatial autocorrelation arises from unobserved spatially correlated factors. However, the primary concern of this thesis is to assess whether the spatial autocorrelation of district-level fertility is due to relationship between fertility rate in a district and the fertility rates of neighbouring districts, capturing the diffusion process explicitly. Therefore, SEM does not explicitly account for the influence of neighbouring fertility rates, which may be crucial in understanding the diffusion process. Therefore, SEM may not adequately capture the spatial autocorrelation of district-level fertility in Ethiopia resulting from the spread of fertility behaviours. Hence, I opt for SLM rather than SEM to study the diffusion process of fertility decline and address spatial autocorrelation in district-level fertility.

3.3.2.3. Objective 5: To explore spatial heterogeneity in relationships between TFRs and both proximate and distal determinants in Ethiopia by using geographically weighted regression between 2000 and 2016.

When the spatial dimension plays a significant role in the relationship between study variables, geographically weighted regression (GWR) can be employed. The GWR model is increasingly utilized in fertility studies as it can identify the factors contributing to fertility fluctuations and capture the spatially heterogeneous relationships between location-specific fertility and (Wang and Chi, 2017, Haque et al., 2019, Obradovic and Vojkovic, 2021).Unlike global linear regression models such as the spatial lag model, the GWR model is a local linear regression approach that can identify specific local trends in the spatial distribution of parameters. In order to examine the spatially varying relationships between study variables, the GWR model conducts separate regressions at each district, which are specified as (Brunsdon et al., 1996, Fotheringham and Oshan, 2016):

$$y(TFR)_j = \beta_0(lon_j, lat_j) + \sum_{m=1}^{N} \beta_m\,(lon_j, lat_j) x_{mj} + \varepsilon_j$$

(lon_j, lat_j) denotes the longitude and latitude coordinates of woreda j. The locations for each district in the GWR analysis is the centroids of each district. The regression coefficients β_{m} are estimated by a geographically weighted matrix of the form:

$$\hat{\beta}(lon_j, lat_j) = [\mathbf{X}^T\mathbf{W}(lon_j, lat_j)\mathbf{X}]^{-1}[\mathbf{X}^T\mathbf{W}(lon_j, lat_j)\mathbf{Y}],$$

where $\hat{\beta}$ is a vector of local estimators of β; $\boldsymbol{X}$ and $\mathbf{Y}$ are vectors of selected explanatory variables and TFRs. $\mathbf{W}(lon_j, lat_j)$ is the diagonal weighting matrix relative to the location of (lon_j, lat_j), which contains 0 in its off-diagonal elements.

$$\mathbf{W}(lon_j, lat_j) = \begin{bmatrix} w_1(lon_j, lat_j) & 0 & 0 \\ 0 & \ldots & 0 \\ 0 & 0 & w_n(lon_j, lat_j) \end{bmatrix}$$

Thus, $\hat{\beta}(lon_j, lat_j)$ varies with the values of $\mathbf{W}(lon_j, lat_j)$. For spatial weights, I adopted Gaussian weights and the bi-square weighting function. The function is as follows:

$$W_{ij} = \begin{cases} [1-(d_{ij}/b)^2]^2, d_{ij} < d_{max} \\ 0 \qquad\qquad , otherwise \end{cases}$$

where d_{ij} is the Euclidean distance between district j for estimation and specific location i for observation. b is the bandwidth size that defines how many neighbouring observations should be included in the weight matrix (Fotheringham, 1997). There are two options for determining the bandwidth: a fixed method and an adaptive method. The fixed method defines a spatial cluster around all regression points using a fixed bandwidth, where the kernel range is determined by the distance to a specific regression point. The adaptive method utilizes various bandwidths to specify spatial clusters surrounding each regression point, as illustrated in Figure 3.6(B). The number of neighbours from a specific regression point determines the kernel range. In cases of sparse data, the kernel range has a larger bandwidth. In this study, I used the adaptive method to determine the bandwidth size due to the uneven distances across Ethiopian districts.

(A) **(B)**

Figure 3.7. Example of (A) fixed and (B) adaptive bandwidth
Source: (Shabrina et al., 2021)

3.4. Closing remarks

Further details of data and methods are covered in Chapter 4 and Chapter 5.

3.5. References

AHETO, J. M. K., & DAGNE, G. A. 2021. Geostatistical analysis, web-based mapping, and environmental determinants of under-5 stunting: evidence from the 2014 Ghana Demographic and Health Survey. *The Lancet Planetary Health*, 5(6), e347-e355.

ALEM, A. Z., YESHAW, Y., KEBEDE, S. A., LIYEW, A. M., TESEMA, G. A., AGEGNEHU, C. D. & TESHALE, A. B. 2020. Spatial Distribution and Determinants of Early Marriage among Married Women in Ethiopia: A spatial and Multilevel Analysis. *BMC Women's Health,* 20**,** 1-13.

ALHO, J. M. 2001. Demographic Techniques: Small-area Estimates and Projections. Neil J. Smelser, Paul B. Baltes (éd.), International Encyclopedia of the Social & Behavioral Sciences, Oxford, Pergamon.

ANDRES, L. A., BHATT, S., DASGUPTA, B., ECHENIQUE, J., GETHING, P. W., ZABLUDOVSKY, G. & JOSEPH, G. 2018. Geo-spatial modeling of access to water and sanitation in Nigeria. *World Bank Policy Research Working Paper*.

ANSELIN, L. 1995. Local indicators of spatial association—LISA. *Geographical analysis,* 27**,** 93-115.

ANSELIN, L. & ARRIBAS-BEL, D. 2013. Spatial fixed effects and spatial dependence in a single cross-section. *Papers in Regional Science,* 92**,** 3-17.

ANSELIN, L. & REY, S. 1991. Properties of tests for spatial dependence in linear regression models. *Geographical analysis,* 23**,** 112-131.

BLANGIARDO, M. & CAMELETTI, M. 2015. *Spatial and spatio-temporal Bayesian models with R-INLA*, John Wiley & Sons.

BLANGIARDO, M., CAMELETTI, M., BAIO, G. & RUE, H. 2013. Spatial and spatio-temporal models with R-INLA. *Spatial and spatio-temporal epidemiology,* 4**,** 33-49.

BONGAARTS, J. 1978. A framework for analyzing the proximate determinants of fertility.

BONGAARTS, J. 1982. The fertility-inhibiting effects of the intermediate fertility variables. *Studies in family planning***,** 179-189.

BONGAARTS, J. 2017. Africa's unique fertility transition. *Population and Development Review,* 43**,** 39-58.

BONGAARTS, J. 2020. Trends in fertility and fertility preferences in sub-Saharan Africa: the roles of education and family planning programs. *Genus,* 76**,** 1-15.

BONGAARTS, J. & WATKINS, S. C. 1996. Social interactions and contemporary fertility transitions. *Population and development review***,** 639-682.

BRUNSDON, C., FOTHERINGHAM, A. S. & CHARLTON, M. E. 1996. Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis,* 28**,** 281-298.

BURGERT-BRUCKER, C. R., J. COLSTON, T. R, and B. ZACHAR. 2013. *Geographic displacement procedure and georeferenced data release policy for the Demographic and Health Surveys*. DHS Spatial Analysis Report No. 7. Calverton, Maryland, USA: ICF International

BURGERT-BRUCKER, C. R., DOMTAMSETTI, T., MARSHALL, A. M. & GETHING, P. 2016a. *Guidance for use of the DHS program modeled map surfaces*, ICF International.

BURGERT-BRUCKER, C. R., DONTAMSETTI, T., MASHALL, A. & GETHING, P. 2016b. Guidance for use of The DHS Program modeled map surfaces. *DHS Spatial Analysis Reports No. 14.* Rockville, Maryland, USA: ICF International.

BURGERT-BRUCKER, C. R., DONTAMSETTI, T. & GETHING, P. W. 2018. The DHS program's modeled surfaces spatial datasets. *Studies in family planning,* 49**,** 87-92.

BURGERT, C. R. 2014. Spatial interpolation with Demographic and Health Survey data: Key considerations. *DHS Spatial Analysis Reports No. 9.* Rockville, Maryland, USA: ICF International.

CASTERLINE, J. B. 2017. Prospects for fertility decline in Africa. *Population and Development Review,* 43**,** 3-18.

CLARK, S., KOSKI, A. & SMITH-GREENAWAY, E. 2017. Recent trends in premarital fertility across sub-Saharan Africa. *Studies in Family Planning,* 48**,** 3-22.
CLELAND, J., BERNSTEIN, S., EZEH, A., FAUNDES, A., GLASIER, A. & INNIS, J. 2006. Family planning: the unfinished agenda. *The Lancet*, 368**,** 1810-1827.
CLELAND, J. & WILSON, C. 1987. Demand theories of the fertility transition: An iconoclastic view. *Population studies,* 41**,** 5-30.
CRESSIE, N. 2015. *Statistics for spatial data*, John Wiley & Sons.
CROFT, T. N., AILEEN M. J. MARSHALL, COURTNEY K. ALLEN, ET AL. 2018. *Guide to DHS Statistics,* Rockville, Maryland, USA: ICF.
CROFT, T. N., MARSHALL, A. M., ALLEN, C. K., ARNOLD, F., ASSAF, S. & BALIAN, S. 2018. Guide to DHS statistics. *Rockville: ICF*, 645.
DIGGLE, P. J., TAWN, J. A. & MOYEED, R. A. 1998. Model-based geostatistics. *Journal of the Royal Statistical Society: Series C (Applied Statistics),* 47**,** 299-350.
DIGGLE, P. J., and ROBEIRO, P, J. 2007. *Model-Based Geostatistics*. 1st ed. Springer Series in Statistics
FINGLETON, B. & LE GALLO, J. 2008. Estimating spatial models with endogenous variables, a spatial lag and spatially dependent disturbances: Finite sample properties. *Papers in Regional Science,* 87**,** 319-339.
FISH, T. D., JANOCHA, B., DONTAMSETTI, T. & MAYALA, B. K. 2020. Geospatial covariates: proxies for mapping urban-related indicators. *DHS Spatial Analysis Reports No. 19.* Rockville, Maryland, USA: ICF.
FOTHERINGHAM, A. S. 1997. Trends in quantitative methods I: stressing the local. *Progress in Human Geography,* 21**,** 88-96.
FOTHERINGHAM, A. S. & OSHAN, T. M. 2016. Geographically weighted regression and multicollinearity: dispelling the myth. *Journal of Geographical Systems,* 18**,** 303-329.
GARCIA, M. & RAJKUMAR, A. S. 2008. *Achieving better service delivery through decentralization in Ethiopia*, World Bank Publications.
GARENNE, M. & JOSEPH, V. 2002. The timing of the fertility transition in sub-Saharan Africa. *World Development*, 30**,** 1835-1843.
GAVRILOVIC, M., PALERMO, T., VALLI, E., VIOLA, F., VINCI, V., HEISSLER, K., RENAULT, M., SERDAN, A. G. G., MUSSA, E. C. & UNICEF 2020. Child Marriage and Ethiopia's Productive Safety Net Program: Analysis of protective pathways in the Amhara region.
GETHING, P., TATEM, A., BIRD, T. & BURGERT, C. 2015. Creating spatial interpolation surfaces with DHS data.
GETHING, P. W. & BURGERT-BRUCKER, C. R. 2017. The DHS Program modeled map surfaces: understanding the utility of Spatial Interpolation for generating indicators at subnational administrative levels. *DHS Spatial Analysis Reports No. 15.* Rockville, Maryland, USA: ICF.
GIORGI, E., FRONTERRê, C., MACHARIA, P. M., ALEGANA, V. A., SNOW, R. W. & DIGGLE, P. J. 2021. Model building and assessment of the impact of covariates for disease prevalence mapping in low-resource settings: to explain and to predict. *Journal of the Royal Society Interface,* 18**,** 20210104.
GOLDSTEIN, J. R. & KLþSENER, S. 2014. Spatial analysis of the causes of fertility decline in Prussia. *Population and Development Review,* 40**,** 497-525.
GRIES, T. & GRUNDMANN, R. 2018. Fertility and modernization: the role of urbanization in developing countries. *Journal of International Development,* 30**,** 493-506.
HAQUE, I., DAS, D. & PATEL, P. P. 2019. Reading the geography of India's district-level fertility differentials: a spatial econometric approach. *Journal of Biosocial Science,* 51**,** 745-774.
HARRIS, R. 2016. *Quantitative geography: The basics*, Sage.
HERTRICH, V. 2017. Trends in age at marriage and the onset of fertility transition in sub-Saharan Africa. *Population and Development Review,* 43**,** 112-137.
HOGAN, D. P. & BIRATU, B. 2004. Social identity and community effects on contraceptive use and intentions in southern Ethiopia. *Studies in Family Planning,* 35**,** 79-90.

HUANG, J., MALONE, B. P., MINASNY, B., MCBRATNEY, A. B. & TRIANTAFILIS, J. 2017. Evaluating a Bayesian modelling approach (INLA-SPDE) for environmental mapping. *Science of the Total Environment,* 609**,** 621-632.
ICF International and Central Statistical Agency. Ethiopia DHS 2000. Addis Ababa, Ethiopia and Calverton, MD: Central Statistical Agency, ICF International, 2000
ICF International and Central Statistical Agency. Ethiopia DHS 2005. Addis Ababa, Ethiopia and Calverton, MD: Central Statistical Agency, ICF International, 2005.
ICF International and Kenya National Bureau of Statistics. Kenya DHS 2014. Nairobi, Kenya and Calverton, MD: Kenya National Bureau of Statistics, ICF International, 2014.
ICF International and Central Statistical Agency. Ethiopia DHS 2016. Addis Ababa, Ethiopia and Calverton, MD: Central Statistical Agency, ICF International, 2016.
JEJEEBHOY, S. J. 1995. Women's education, autonomy, and reproductive behaviour: Experience from developing countries. *OUP Catalogue*.
KEBEDE, E., GOUJON, A. & LUTZ, W. 2019. Stalls in Africa's fertility decline partly result from disruptions in female education. *Proceedings of the National Academy of Sciences,* 116**,** 2891-2896.
KRAINSKI, E. T., GõMEZ-RUBIO, V., BAKKA, H., LENZI, A., CASTRO-CAMILO, D., SIMPSON, D., LINDGREN, F. & RUE, H. 2018. *Advanced spatial modeling with stochastic partial differential equations using R and INLA*, CRC Press.
LAELAGO, T., HABTU, Y. & YOHANNES, S. 2019. Proximate determinants of fertility in Ethiopia; an application of revised Bongaarts model. *Reproductive health,* 16**,** 13.
LAKEW, Y., REDA, A. A., TAMENE, H., BENEDICT, S. & DERIBE, K. 2013. Geographical variation and factors influencing modern contraceptive use among married women in Ethiopia: evidence from a national population based survey. *Reproductive health,* 10**,** 1-10.
LERCH, M. 2019. Regional variations in the rural-urban fertility gradient in the global South. *PloS one,* 14**,** e0219624.
LINDGREN, F. & RUE, H. 2015. Bayesian spatial modelling with R-INLA. *Journal of Statistical Software,* 63**,** 1-25.
LINDGREN, F., RUE, H. & LINDSTRøM, J. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 73**,** 423-498.
LINDSTROM, D. P., KIROS, G.-E. & HOGAN, D. P. 2009. Transition into first intercourse, marriage, and childbearing among Ethiopian women. *Genus,* 65**,** 45.
MAY, J. F. & ROTENBERG, S. 2020. A Call for Better Integrated Policies to Accelerate the Fertility Decline in Sub-Saharan Africa. *Studies in Family Planning,* 51**,** 193-204.
MAYALA, B. K., BHATT, S. & GETHING, P. 2020. Predicting HIV/AIDS at subnational levels using DHS covariates related to HIV. *DHS Spatial Analysis Reports No. 18.* Rockville, Maryland, USA: ICF.
MAYALA, B. K., DONTAMSETTI, T., FISH, T. D. & CROFT, T. 2019a. *Interpolation of DHS survey data at subnational administrative Level 2,* DHS Program, ICF.
MAYALA, B. K., DONTAMSETTI, T., FISH, T. D. & CROFT, T. N. 2019b. Interpolation of DHS survey data at subnational administrative Level 2. *DHS Spatial Analysis Reports No. 17.* Rockville, Maryland, USA: ICF.
MEKONNEN, Y., TELAKE, D. S. & WOLDE, E. 2018. Adolescent childbearing trends and sub-national variations in Ethiopia: a pooled analysis of data from six surveys. *BMC pregnancy and childbirth*, 18, 1-13.
MERCER, L. D., WAKEFIELD, J., PANTAZIS, A., LUTAMBI, A. M., MASANJA, H., & CLARK, S. (2015). Space-time smoothing of complex survey data: small area estimation for child mortality. *The annals of applied statistics*, 9(4), 1889.
MONTGOMERY, M. R. & CASTERLINE, J. B. 1993. The diffusion of fertility control in Taiwan: Evidence from pooled cross-section time-series models. *Population studies,* 47**,** 457-479.
MORAGA, P. 2019. *Geospatial Health Data: Modeling and Visualization with R-INLA and Shiny*, Chapman and Hall/CRC.

MURTIN, F. 2013. Long-term determinants of the demographic transition, 1870–2000. *Review of Economics and Statistics,* 95**,** 617-631.
OBRADOVIC, S. L. & VOJKOVIC, G. 2021. The Impact of Spatial-Demographic Disparities on Fertility Variations in Serbian Municipalities—An Example of Hidden Content Research. *Hidden Geographies***,** 379.
OUADAH-BEDIDI, Z., VALLIN, J. & BOUCHOUCHA, I. 2012. Unexpected developments in Maghrebian fertility. *Population & Sociétés*.
PEREZ-HAYDRICH, C., WARREN, J. L., BURGERT, C. R. & EMCH, M. E. 2013. Guidelines on the use of DHS GPS data. *DHS Spatial Analysis Reports No. 8.* Calverton, Maryland, USA: ICF International.
POGGIO, L., GIMONA, A., SPEZIA, L. & BREWER, M. J. 2016. Bayesian spatial modelling of soil properties and their uncertainty: The example of soil organic matter in Scotland using R-INLA. *Geoderma,* 277**,** 69-82.
REINER JR, R. C., WIENS, K. E., DESHPANDE, A., BAUMANN, M. M., LINDSTEDT, P. A., BLACKER, B. F., TROEGER, C. E., EARL, L., MUNRO, S. B. & ABATE, D. 2020. Mapping geographical inequalities in childhood diarrhoeal morbidity and mortality in low-income and middle-income countries, 2000–17: analysis for the Global Burden of Disease Study 2017. *the lancet,* 395**,** 1779-1801.
ROBINSON, W. C. 1963. Urbanization and fertility: the non-western experience. *The Milbank Memorial Fund Quarterly,* 41**,** 291-308.
RUE, H., MARTINO, S. & CHOPIN, N. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology),* 71**,** 319-392.
SALVATI, L., BENASSI, F., MICCOLI, S., RABIEI-DASTJERDI, H. & MATTHEWS, S. A. 2020. Spatial variability of total fertility rate and crude birth rate in a low-fertility country: patterns and trends in regional and local scale heterogeneity across Italy, 2002–2018. *Applied Geography,* 124**,** 102321.
SCHMERTMANN, C. P., CAVENAGHI, S. M., ASSUNCAO, R. M., & POTTER, J. E. (2013). Bayes plus Brass: estimating total fertility for many small areas from sparse census data. *Population studies*, *67*(3), 255-273.
SHABRINA, Z., BUYUKLIEVA, B. & NG, M. K. M. 2021. Short-Term Rental Platform in the Urban Tourism Context: A Geographically Weighted Regression (GWR) and a Multiscale GWR (MGWR) Approaches. *Geographical Analysis,* 53**,** 686-707.
SHAPIRO, D. 2012. Women's education and fertility transition in sub-Saharan Africa. *Vienna yearbook of population research***,** 9-30.
SHAPIRO, D. & GEBRESELASSIE, T. 2014. Marriage in sub-Saharan Africa: Trends, determinants, and consequences. *Population Research and Policy Review,* 33**,** 229-255.
SHAPIRO, D. & TENIKUE, M. 2017. Women’s education, infant and child mortality, and fertility decline in urban and rural sub-Saharan Africa. *Demographic Research,* 37**,** 669-708.
SINGH, A., KUMAR, K., PATHAK, P. K., CHAUHAN, R. K. & BANERJEE, A. 2017. Spatial patterns and determinants of fertility in India. *Population,* 72**,** 505-526.
SMITH-GREENAWAY, E. & CLARK, S. 2018. Women's marriage behavior following a premarital birth in sub-Saharan Africa. *Journal of Marriage and Family,* 80**,** 256-270.
SOBOTKA, T. 2017. Post-transitional fertility: the role of childbearing postponement in fuelling the shift to low and unstable fertility levels. *Journal of biosocial science,* 49**,** S20-S45.
TEGEGNE, T. K., CHOJENTA, C., FORDER, P. M., GETACHEW, T., SMITH, R. & LOXTON, D. 2020. Spatial variations and associated factors of modern contraceptive use in Ethiopia: a spatial and multilevel analysis. *BMJ open,* 10**,** e037532.
TOWRISS, C. A. & TIMèUS, I. M. 2018. Modelling period fertility: Schooling and intervals following a birth in Eastern Africa. *Population studies,* 72**,** 75-90.
UNOCHA. 2020. *Ethiopia: Administrative map* [Online]. UNOCHA. Available: https://reliefweb.int/map/ethiopia/ethiopia-administrative-map-october-2020 [Accessed October 2020].

VITALI, A. & BILLARI, F. C. 2017. Changing determinants of low fertility and diffusion: A spatial analysis for Italy. *Population, Space and Place,* 23**,** e1998.
WANG, D. & CHI, G. 2017. Different Places, Different Stories: A Study of Spatial Heterogeneity of County-Level Fertility in China. *Demographic research,* 37**,** 493.
WATKINS, S. C. The fertility transition: Europe and the Third World compared. Sociological forum, 1987. Springer, 645-673.
WATKINS, S. C. 1990. From local to national communities: The transformation of demographic regimes in Western Europe, 1870-1960. *Population and development review***,** 241-272.
WHITE, M. J. 1986. Segregation and diversity measures in population distribution. *Population index***,** 198-221.
WILSON, K. & WAKEFIELD, J. 2020. Pointless spatial modeling. *Biostatistics,* 21**,** e17-e32.

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1704116	**Title**	Mr
First Name(s)	Myunggu		
Surname/Family Name	Jung		
Thesis Title	Spatial aspects of fertility change in Ethiopia between 2000 and 2016: a district-level analysis		
Primary Supervisor	Kazuyo Machiyama		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	BMJ Global Health
Please list the paper's authors in the intended authorship order:	Myunggu Jung, Christopher I Jarvis, Ian M Timæus, Kazuyo Machiyama
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I conceived of study design and the statistical analysis plan, conducted the statistical analysis and wrote the first draft with feedback from co-authors.

SECTION E

Student Signature	[redacted]
Date	20/12/2022

Supervisor Signature	[redacted]
Date	20/12/2022

<u>Chapter 4</u>

Paper 1: Geographical distribution of fertility in Ethiopia between 2000 and 2016: a district-level analysis

Paper 1: Geographical distribution of fertility in Ethiopia between 2000 and 2016: a district-level analysis

Myunggu Jung[1], Christopher I Jarvis[2], Ian M Timæus[1,3], Kazuyo Machiyama[1]

1. Department of Population Health, London School of Hygiene & Tropical Medicine, London, UK
2. Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK
3. Centre for Actuarial Research, University of Cape Town, Cape Town, South Africa

4. Chapter 4: Geographical distribution of fertility in Ethiopia between 2000 and 2016: a district-level analysis

4.1. Overview

In Chapter 4, I applied Bayesian model-based geostatistics (MBG) to generate maps of fertility and key selected proximate and distal determinants in 981 districts in Ethiopia between 2000 and 2016. These maps have policy implications for the decentralised health service system at the district level in Ethiopia. This paper describes and explores the changes in sub-national variations in fertility and key proximate and distal determinants between 2000 and 2016.

Objective 1	To estimate TFRs and key selected proximate and distal determinants for 981 districts in 2000, 2005, 2011 and 2016 by using a geostatistical modelling approach.
Objective 2	To describe and explore spatial and temporal patterns of TFR and key selected proximate and distal determinants at the district level in 2000, 2005, 2011 and 2016.

4.2. Role of candidate

I conceived of the study design and statistical analysis plan which was agreed by the co-authors. I conducted the statistical analysis and wrote the first draft of the manuscript with feedback and inputs provided from Christopher I Jarvis, Ian M Timæus, and Kazuyo Machiyama.

4.3. Abstract

Background

Understanding and acknowledging fertility differences between districts in Ethiopia is important because districts, known as woredas, serve as vital administrative units for health policy and planning. While previous studies have often examined geographical variations in fertility between urban and rural areas or at the regional level, limited information is available regarding geographical variations in fertility at the district level due to the scarcity of district-level data in Ethiopia. Therefore, the objective of this study is to describe and explore geographical variations in fertility rates and key proximate and distal determinants at the district level in Ethiopia between 2000 and 2016.

Methods

This cross-sectional study analysed data from the 2000, 2005, 2011, and 2016 Ethiopia Demographic and Health Surveys. It employed a Bayesian model-based geostatistical approach using a stochastic partial differential equation (SPDE) within the integrated nested Laplace approximation (INLA) framework. The study aimed to estimate various indicators, including the total fertility rate (TFR), modern contraceptive prevalence (mCP), median age at first marriage, proportion of women living in urban areas, proportion of women with secondary or higher education, and ethnolinguistic diversity for 981 districts between 2000 and 2016. Mapping analysis was conducted to investigate changes in the spatial patterns of these variables at the district level in Ethiopia.

Results

In 2000 and 2005, geographical variations in fertility at both the regional and district levels in Ethiopia were relatively small. However, district-level fertility diverged both between and within regional states in 2011 and 2016, resulting in wider variations within the same regions. Notably, distinct spatial patterns of fertility have emerged, with lower fertility rates gradually spreading from the capital to the northern and western parts of Ethiopia. These spatial patterns of district-level TFR were partly associated with the locations of urban areas but primarily driven by changes in the spatial patterns of mCP.

Conclusion

Although district-level fertility did not vary much between districts in 2000 and 2005, substantial variations in district-level fertility have emerged in recent years, even within the same region, in Ethiopia. This result implies that focusing solely on national or regional data provides an inadequate description of the geographical variations in contemporary fertility in Ethiopia. This paper contributes to our understanding of changes in the spatial pattern of fertility and the key factors determining fertility at the district level in Ethiopia.

4.4. Introduction

The Sustainable Development Goals (SDGs), specifically target 17.18, aim to ensure that no one is left behind by improving access to timely and accurate data disaggregated by geographic location and other relevant factors within national contexts by 2030 (UN, 2015). The growing availability of georeferenced data in low-income countries, including data collected through the Demographic and Health Survey (DHS) programme, has facilitated the use of geostatistics to generate high-resolution maps of health and demographic indicators in sub-Saharan Africa (SSA).

Understanding and acknowledging subnational variations in fertility and its determinants at the district level is crucial for effective public health policy planning in SSA countries. Many SSA countries have implemented decentralisation of healthcare delivery to enhance the quality of local health services (Zon et al., 2017). Recent research has demonstrated the significance of high-resolution spatial estimates in mapping demographic and health indicators at the subnational level in SSA countries, revealing variations that may be masked by aggregate or national-level indicators (Hosseinpoor et al., 2018, Dwyer-Lindgren et al., 2019, Burke et al., 2016, Reiner Jr et al., 2020, Bhattacharjee et al., 2019, Graetz et al., 2018, Osgood-Zimmerman et al., 2018). While administrative boundaries are commonly used for decision-making by national policymakers and program managers, accurate high-resolution spatial estimates of health indicators can offer valuable tools for spatially targeted health interventions. Hence, estimates that align with administrative boundaries would be more beneficial for national policymakers engaged in subnational planning. Although recent studies have explored fertility and mortality at the first administrative unit (Admin 1), such as states or provinces, in SSA (Pezzulo et al., 2021, Li et al., 2019b), there is a need for estimates at even lower administrative levels, such as Admin 2 or 3, as health program implementation is often decentralized to those levels (Mayala et al., 2019a, Janocha et al., 2021).

In Ethiopia, the district level, known as 'woreda' in Amharic, plays a crucial role as an administrative unit for health policy and implementation. The country's health system follows a decentralised approach, empowering district administrations with decision-making authority. This is exemplified by the National Health Sector Transformation Plan II (HSTP II) agenda, which emphasizes "Woreda-Based Health Sector Planning (WBHSP)." The aim of WBHSP is to mainstream the Sustainable Development Goals (SDGs) at the district level, ensuring that no one is left behind by customizing national health programs to the specific context of each district in Ethiopia (MoH, 2021). The adoption of WBHSP initially aimed to address health disparities between districts in response to significant geographical variations in fertility and health outcomes in Ethiopia. For instance, in 2016, the total fertility rate (TFR) was 1.8 in Addis Ababa and 7.2 in the Somali regional state (MoH, 2015, ICF, 2016). Ethiopia stands as the second most populous nation in sub-Saharan Africa (SSA), with an estimated population of about 115 million in 2021. The country's population is highly diverse, encompassing over 90 different ethno-linguistic groups. Ethno-linguistic identity serves as a key criterion in determining sub-national administrative boundaries (Abbink, 2011). The eleven regions are divided along ethno-linguistic lines, comprising five regional states dominated by a single ethno-linguistic group (Amhara, Tigray, Afar, Oromia, and Somali regional states), four multi-ethnic regional states (Harari, Southern Nation, Nationalities and Peoples (SNNP), Benishangul-Gumuz, Gambella regional states), and two multi-ethnic cities (Addis Ababa and Dire-Dawa) (Figure 4.1). Additionally, two eastern regional states (Afar and Somali) and two western regional states (Benishangul-Gumuz and Gambella) are often categorized as the Developing Regional States (DRS). These regions are primarily inhabited by pastoral communities, whose health and socio-economic outcomes often lag behind the other regions in Ethiopia (Chekole et al., 2019, Getnet et al., 2017, Alemayehu et al., 2016).

Ethiopia's geography is therefore often characterized by pervasive ethnolinguistic, health, demographic, and socioeconomic heterogeneities. In the realm of demographic theory, geographical variations in fertility are often attributed to either different socioeconomic conditions (the adaptationist approach) or the spatial diffusion of new information and the social acceptability of fertility control (the diffusionist approach). Recent studies on fertility in high- and middle-income countries (HMICs) have revealed marked geographic variations in fertility at the district level and assessed how adaption and diffusion effects can jointly account for sub-national fertility variations (Campisi et al., 2020, Wang and Chi, 2017, Vitali and Billari, 2017, Sabater and Graham, 2019, Haque et al., 2019). Therefore, both theoretical arguments and empirical evidence suggest that geographical variations in fertility at the district level in Ethiopia matter. However, very little is known about geographical variations in fertility at the district level due to the shortage of district-level data in Ethiopia.

In this paper, I aim to examine the changes in spatial patterns of fertility at the district level in Ethiopia, considering both its proximate and distal determinants. Proximate determinants include biological and behavioural factors that directly influence fertility rates, while distal determinants encompass socioeconomic and cultural factors that indirectly affect fertility rates through the proximate determinants. To achieve this, I estimate fertility rates and the key proximate and distal determinants using data from the Ethiopia Demographic and Health Surveys (EDHS) conducted between 2000 and 2016. Our approach involves utilising Bayesian model-based geostatistics (MBG) to generate maps illustrating the geographic distribution of fertility and its determinants at the district level. No prior study has specifically explored the dynamics of geographical variations in fertility at the district level in Ethiopia. Therefore, the present paper intends to fill this research gap by offering additional insights into the geographic variations in district-level fertility across Ethiopia.

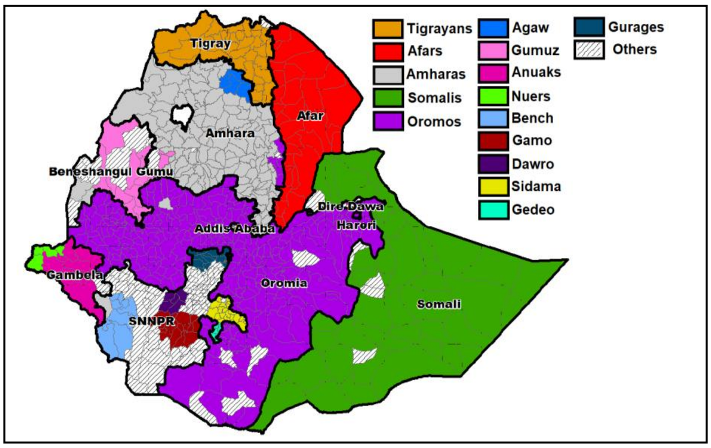


Figure 4.1. Majority ethnic group in each district of Ethiopia according to the 2007 Census
Data source: The 2007 Ethiopia Population and Housing Census

4.5. Methods

4.5.1. Data

I obtained data from EDHS conducted in 2000, 2005, 2011, 2016 (ICF, 2000, ICF, 2005, ICF, 2011, ICF, 2016) and restricted the sample to women aged between 15-49. The EDHS samples are designed to provide estimates for a range of demographic and health indicators that are nationally representative, as well as representative for the eleven regional states. The sampling strategy of the EDHS involves stratification by regional state and urban/rural areas. The samples are selected in two stages. In the first stage, primary sampling units (PSUs) are chosen with a probability proportional to their size within each stratum. In the second stage, approximately 24-35 households are randomly selected from each PSU. The EDHS also provides GPS coordinates for the PSUs, which allows for spatial analysis. To ensure respondent confidentiality, these geographical coordinates are displaced by up to 2 km in urban areas and up to 5 km in rural regions. In our study, we excluded 6, 9, 27, and 21 PSUs without GPS coordinates from the 2000, 2005, 2011, and 2016 EDHS datasets, respectively. As a result, our final analysis included 533, 526, 569, and 622 PSUs from the respective survey years (see Appendix 2). To map the distribution of the total fertility rate (TFR) and key proximate and distal determinants, I obtained corresponding district polygons from UNOCHA Ethiopia (UNOCHA, 2020). These polygons cover 11 regional states, 90 zones, and 981 district-level boundaries. Ethical approval for our study was obtained from the London School of Hygiene & Tropical Medicine, UK, Ethics Committee (25580).

4.5.2. Measurement of study variables

I undertook a subnational analysis of cross-sectional surveys collected from EDHS between 2000 and 2016 to describe changes in spatial patterns of total fertility and two proximate and three distal determinants across 981 districts (Admin 3) in Ethiopia. To calculate the study variables, we aggregated the data at the primary sampling unit (PSU) level, taking into account survey-specific weights as specified in the DHS manuals. It is crucial to acknowledge the impact of geographical displacement when conducting spatial modelling using DHS surveys, as it affects the precise location of individual data points. The EDHS data incorporates buffers of 2 km for urban PSUs and 5 km for rural PSUs around the recorded geographical coordinates to address this displacement.

I measured fertility by means of subnational TFRs calculated using methods described by the DHS Guide to Statistics (Croft et al., 2018). In the DHS, the TFR is estimated as the average number of livebirths a women would have if she was subject to the age-specific fertility rates estimated from the number of live births occurred during the three years preceding the survey through her reproductive years (15-49 years).

Among the four proximate determinants of fertility (contraception, marriage, abortion, postpartum infecundability) (Bongaarts, 2015), I restricted our study to include two proximate determinants: a) contraception and b) marriage. We did not consider the other two determinants because a recent study found that the differences in the inhibiting effects of contraception and marriage on fertility between the eleven regional states in 2011 and 2016 were much larger than those of the other two determinants in Ethiopia (Laelago et al., 2019). To measure the use of contraception, we employed the modern contraceptive prevalence (mCP), which is defined as the proportion of currently married women currently using any modern method of contraception. As an indicator of nuptiality, we used the median age at first marriage, which is defined as the median age in years when women first start living with their spouse. Childbearing outside of marriage is uncommon in Ethiopia, and the age at first marriage is therefore an important proximate determinant influencing regional variations in fertility in the country (Gurmu and Etana, 2014).

Although there is ongoing debate regarding the most important distal determinants of fertility, our analysis focused on three specific distal determinants: a) the proportion of women living in urban areas, b) the proportion of women with secondary or higher education, and c) ethnolinguistic diversity at the zonal level. We selected these determinants because Ethiopia exhibited the largest fertility differentials between residential areas (rural and urban) and female education levels (secondary and lower than secondary) among all the Sub-Saharan African countries surveyed by the DHS program (see Appendix 2). To ensure comparability across different survey periods, we standardized the definition of secondary education as nine years of schooling. This adjustment was necessary because the definition of secondary education changed from seven years in the 2000 and 2005 EDHS surveys to nine years in the 2011 and 2016 EDHS surveys.

In addition, previous studies have provided evidence that geographical areas sharing common culture and language tend to experience fertility decline around the same time, regardless of their socioeconomic status. (Watkins, 1987, Watkins, 1990, Bongaarts and Watkins, 1996). To measure the similarity or diversity in ethno-languages between neighbouring districts in Ethiopia, we used the standardised index of diversity at the zonal level. This index, commonly known as the 'entropy index,' has been widely employed to assess the impact of ethnic or cultural diversity on fertility behaviour (Hogan and Biratu, 2004, De Broe and Hinde, 2006). The index of ethnolinguistic diversity, as defined by White (1986), is calculated as follows (White, 1986);

$$ID_z = -\sum_{g=1}^{g=G} dv_{zg} \ln(dv_{zg})$$

Where $z = 1$ to 90 (The number of zonal areas for this study), and g refers to the number of ethnolinguistic groups. Therefore,

$$dv_{zg} = \frac{N_{zg}}{N_z}$$

where N_{zg} is the number of persons in the g^{th} ethnic group in the z^{th} local area, N_z denotes the total population size of the z^{th} local area, and G refers to the total number of ethnic groups in z^{th} local area. Lower values of the entropy index (close to zero) indicate a higher degree of similarity in the ethnolinguistic composition within the zonal areas, while higher values of the index signify greater ethnolinguistic diversity within those areas.

4.5.3. Bayesian Model based Geostatistics with INLA-SPDE

To estimate TFR and key proximate and distal determinants at the district level, we followed a two-step process. The first step involved estimating the high-resolution spatial distribution of the study variables. In the second step, we aggregated the estimates from the model surface to the spatial polygons representing the districts.

In the first step, I employed Bayesian model-based geostatistics (MBG) using a stochastic partial differential equation (SPDE) approach in the integrated nested Laplace approximations (INLA) to predict TFRs and key selected proximate and distal determinants at a continuous spatial resolution. Detailed information about the theory and implementation of INLA-SPDE can be found elsewhere (Lindgren and Rue, 2015). To put it briefly, using INLA for Bayesian inference offers benefits over Markov Chain Monte Carlo techniques, which often face issues with dense covariate matrices that increase computation processing times. INLA provides faster and more efficient estimates of the posterior marginal distribution. Given that spatial processes are often described by a Gaussian random field with Matérn covariance functions, the INLA-SPDE technique is suitable for spatial interpolation (Krainski et al., 2018). Moreover, the MBG technique was recommended by the DHS Spatial Interpolation Working Group as one of the most suitable methods for producing interpolated surfaces. Point estimates such as mean and standard deviation of estimations can be presented as distributions of estimates (or posterior estimates) generated by the Bayesian framework. This approach is useful when modelling DHS data at a smaller geographical scale than what the DHS was designed to represent. The INLA-SPDE was carried out using the R package R-INLA implemented in the R software. Similar modelling frameworks using the Bayesian INLA-SPDE model have been employed in previous research using DHS data to provide high-resolution maps of health indicators (Mayala et al., 2019b, Fish et al., 2020). For three of the variables, TFR and median age at first marriage, ethnolinguistic diversity, conditional on the true mean value μ_i at location i=1…, n, we assumed that the variable (Y_i) follows a Gaussian distribution, $Y_i \sim Normal(\mu_i, \sigma^2)$ and μ_i=$\beta_0 + u(S_i) + \varepsilon_i$. Here, μ_i equal to the sum of an intercept, β_0, and a spatially structured random effect $u(S_i)$ which is a zero-mean Gaussian process with Matérn covariance function. For the three other variables, mCP, proportion of women living in urban areas and proportion of women with secondary or higher education, conditional on the true proportion π_i at location *i=1…, n,* the number of positive outcomes Y_i out of N_i people sampled follows a binomial distribution, $Y_i \sim Binomial(N_i, \pi_i)$ and $logit(\pi_i)$= $\beta_0 + u(S_i) + \varepsilon_i$. To check for the evidence of spatial correlation and justify the introduction of the $u(S_i)$, I compared the observed variograms to a 95% pointwise envelope based on 1,000 Monte Carlo simulations (Diggle and Ribeiro, 2007).

Several studies have included geospatial covariates as fixed effects, such as land surface temperature and average monthly rainfall for disease prevalence or environmental mapping (Giorgi et al., 2021, Huang et al., 2017, Reiner Jr et al., 2020). Although these covariates are likely associated with environmental variables and diseases, they are less likely to be associated with fertility and the key selected determinants of fertility. Furthermore, the aim of this study is to describe geographical variations in fertility rather than investigating, for instance, the changes in the estimation of study variables at the district level associated with a one unit increase in the covariates. Therefore, we considered β_0 as fixed effect and $u(S_i)$ as spatial random effect in the model. We then predicted values of study variables at 145,978 surface pixels, which is approximately equivalent to a 3 km² grid in Ethiopia. These models were fit independently to the data for each year, resulting in modelled outcomes in 2000, 2005, 2010, and 2016. To assess uncertainty, we represented it as the width of the 95% credible interval, a method frequently employed by the DHS program to measure associated uncertainty in modelled surfaces (Janocha et al., 2021, Burgert-Brucker et al., 2018). I further conducted a comparison between the observed estimates from EDHS and the modelled predictions at the regional level to evaluate the goodness-of-fit.

In the second step, I used a simple mean approach to aggregate the point estimate model surface pixels, allowing us to calculate the levels of Total Fertility Rate (TFR) and its key proximate and distal determinants for each of the 981 districts in Ethiopia between 2000 and 2016.

4.6. Results

4.6.1. Model fitness at the regional level and variations between regional states

I compared the observed variogram of the geostatistical data of study variables at the EDHS PSU level to a 95% pointwise envelope. I found that there is evidence of spatial correlation since the observed variogram lies partly outside the 95% pointwise envelope (Appendix 6). Table 4.1 shows the parameter estimates (Intercept (β_o), variances of the Gaussian process (σ_x^2) and the nugget effect (σ_e^2), and nominal range) with the 95% credible intervals for the INLA-SPDE model. It also shows that there are high spatial correlations of study variables, because the estimate for the variance of the spatially structured effect σ_x^2 is generally higher than the unstructured effect σ_e^2. Furthermore, for TFR, mCP, and the proportions of women living in urban areas and with secondary education, the practical range, which indicates the distance at which spatial correlation between observations becomes negligible, expanded between 2000 and 2016. This implies that the spatial correlation between these variables strengthened over longer distances during the periods. On the other hand, the practical range for median age at first marriage decreased and remained relatively stable for the Index of ethnolinguistic diversity.

Table 4.1. Parameter estimates and corresponding 95% confidence intervals (CI) of the INLA-SPDE model

	2000	2005	2011	2016
Total fertility rate	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)
Intercept (β_o)	5.57 (5.37, 5.78)	5.38 (5.13, 5.63)	5.31(5.06, 5.67)	4.91 (4.64, 5.18)
σ_x^2	3.44 (2.79, 4.24)	3.91 (2.98, 5.13)	4.90 (4.05, 5.93)	3.58 (2.82, 4.54)
σ_e^2	1.08 (0.51, 2.28)	2.12 (0.91, 4.93)	1.12 (0.62, 2.03)	1.19 (0.97, 3.75)
Practical range(km)	10.77 (7.56, 15.33)	13.53 (8.67, 21.09)	10.47 (22.71, 43.41)	25.50 (19.08, 41.25)
Modern contraceptive prevalence				
Intercept (β_o)	0.06 (0.05, 0.08)	0.14 (0.12, 0.15)	0.25 (0.15, 0.35)	0.34 (0.25, 0.43)
σ_x^2	0.01 (0.00, 0.02)	0.03 (0.01, 0.04)	0.04 (0.02, 0.06)	0.09 (0.05, 0.12)
σ_e^2	0.01 (0.00, 0.02)	0.02 (0.01, 0.04)	0.02 (0.01, 0.03)	0.05 (0.02, 0.08)
Practical range(km)	12.93 (8.67, 19.23)	6.09 (18.27, 31.50)	89.52 (52.80, 181.77)	216.66 (139.77, 275.88)
Median age at first marriage				
Intercept (β_o)	16.16 (15.24, 17.07)	16.23 (5.30, 17.17)	16.54 (15.84, 17.23)	16.95 (16.62, 17.27)
σ_x^2	1.78 (0.95, 3.31)	1.48 (0.71, 3.11)	1.32 (0.75, 2.32)	0.89 (0.61, 1.30)
σ_e^2	1.07 (0.31, 3.81)	1.35 (0.31, 5.94)	1.59 (0.51, 4.99)	1.51 (0.68, 3.38)
Practical range(km)	367.44 (234.51, 575.49)	410.66 (235.17, 586.15)	304.23 (187.74, 493.08)	194.61 (71.94, 246.12)
Proportion of women living in urban areas				
Intercept (β_o)	0.07 (5.37, 5.78)	0.08 (5.13, 5.63)	0.12 (5.06, 5.67)	0. 17 (4.64, 5.18)
σ_x^2	0.01 (0.00, 0.02)	0.02 (0.01, 0.04)	0.03 (0.01, 0.04)	0.05 (0.03, 0.08)
σ_e^2	0.01 (0.00, 0.02)	0.01 (0.00, 0.02)	0.02 (0.01.0.03)	0.04 (0.02.0.06)
Practical range(km)	16.74 (11.82, 21.69)	19.62 (12.82, 26.58)	24.92 (22.47, 27.39)	36.36 (33.09, 39.63)
Proportion of women with secondary or higher education				
Intercept (β_o)	0.08 (5.37, 5.78)	0.10 (5.13, 5.63)	0.13 (0.05, 0.09)	0.18 (0.10, 0.14)
σ_x^2	0.01 (0.00, 0.02)	0.01 (0.00, 0.02)	0.03 (0.01, 0.04)	0.06 (0.04, 0.08)
σ_e^2	0.01 (0.00, 0.02)	0.01 (0.00, 0.02)	0.02 (0.01.0.03)	0.04 (0.02.0.05)
Practical range(km)	14.94 (11.37, 19.68)	14.19 (10.68, 18.87)	16.59 (11.07, 22.89)	18.93 (13.32, 26.88)
Index of ethnolinguistic diversity				
Intercept (β_o)	0.57 (0.44, 0.70)	0.59 (0.42, 0.75)	0.59 (0.41, 0.70))	0.52 (0.38, 0.65)
σ_x^2	**0.19(0.15 ,0.26)**	**0.18 (0.13, 0.25)**	**0.17(0.13, 0.24)**	**0.19 (0.14, 0.26)**
σ_e^2	0.01 (0.00,0.02)	0.01 (0.00, 0.03)	0.01(0.00. 0.03)	0.01(0.00. 0.03)
Practical range(km)	146.76 (126.81, 170.82)	156.08 (120.32, 201.62)	157.32 (124.74, 198.42)	145.02 (115.47, 182.16)

Note: σ_x^2 = variance of the Gaussian process, σ_e^2= Variance of the nugget effect.

The modelled estimates generally correspond to the estimates observed from EDHS at the regional level, although some small under- and over-estimation was observed, for instance, for TFR in Addis Ababa and Somali regional state in 2016 (Figure 4.2). Therefore, we used the predicted values for further analysis.

In 2000 and 2005, the TFRs of the eight regional states, excluding the three urban regional states (Addis Ababa, Dire-Dawa, and Harari), were relatively similar, around 5. However, in 2011 and 2016, the regional-level TFRs diverged. Some regional states, such as Oromia, Somali, and Afar, still had TFRs around 5 or even experienced an increase in fertility, while other regional states showed a decline in fertility (Figure 4.2A).

Similarly, the mCPs in the eight regional states, excluding the three urban regional states (Addis Ababa, Dire-Dawa, and Harari), were very low and clustered under 20% in 2000 and 2005. However, in 2011 and 2016, the regional-level mCPs diverged. The mCPs of Somali and Afar regional states still remained under 20%, whereas other regional states, including Addis Ababa, experienced increases in mCP in 2011 and 2016 (Figure 4.2B). On the other hand, between 2000 and 2011, the median age at first marriage differed between regional states. However, by 2016, the median age at first marriage of the eleven regional states converged around 17 years old, as most regional states experienced increases in median age at first marriage between 2011 and 2016. Notably, larger increases in the median age at first marriage, almost 2 years, were observed in Addis Ababa and Amhara regional state between 2011 and 2016 compared to other regions (4.2C)

The three urban regional states (Addis Ababa, Dire-Dawa, and Harari) exhibited higher proportions of urban populations and women with secondary or higher education compared to the other eight regional states between 2000 and 2016 (Figure 4.2D and 4.2E). Regional proportions of these two distal determinants remained relatively similar from 2000 to 2011, but started to diverge by 2016, especially in the case of women with secondary or higher education. Ethnolinguistic diversity remained consistent throughout the entire period from 2000 to 2016, but the three urban regional states and the Gambela and Benisangul-Gumz regions displayed greater ethnolinguistic diversity at the zonal level compared to other regions (Figure 4.2F).

4.6.2. Variations within regional states

In addition to regional variations in study variables, sub-regional variations were also observed (Figure 4.3). In particular, between 2000 and 2005, the eight regional states other than the three urban regional states (Addis Ababa, Dire-Dawa and Harari) exhibited relatively similar median levels of TFRs and mCPs, with small variations within the same regional states (Figure 4.3A and 4.3B). However, in 2011 and 2016, TFRs and mCPs diverged both between and within regional states, implying that variations in TFRs and mCPs in districts within the same regional state widened recently. While there were significant sub-regional variations in ethnolinguistic diversity within each regional state, these variations remained relatively stable between 2000 and 2016 (Figure 4.3 F). It is worth noting that although the SNNP regional state is the most ethnolinguistically diverse region in Ethiopia, comprising more than 45 ethnolinguistic groups, the level of ethnolinguistic diversity at the zone level was relatively low. Sub-regional variations in the proportions of women living in urban areas within each regional state were small in 2000 and 2016, except for the three urban regional states (Addis Ababa, Dire-Dawa, and Harari) (Figure 4.3D). While sub-regional variations in the proportions of women with secondary or higher education within each regional state were small until 2011, they became more pronounced in most regional states by 2016 (Figure 4.3E).

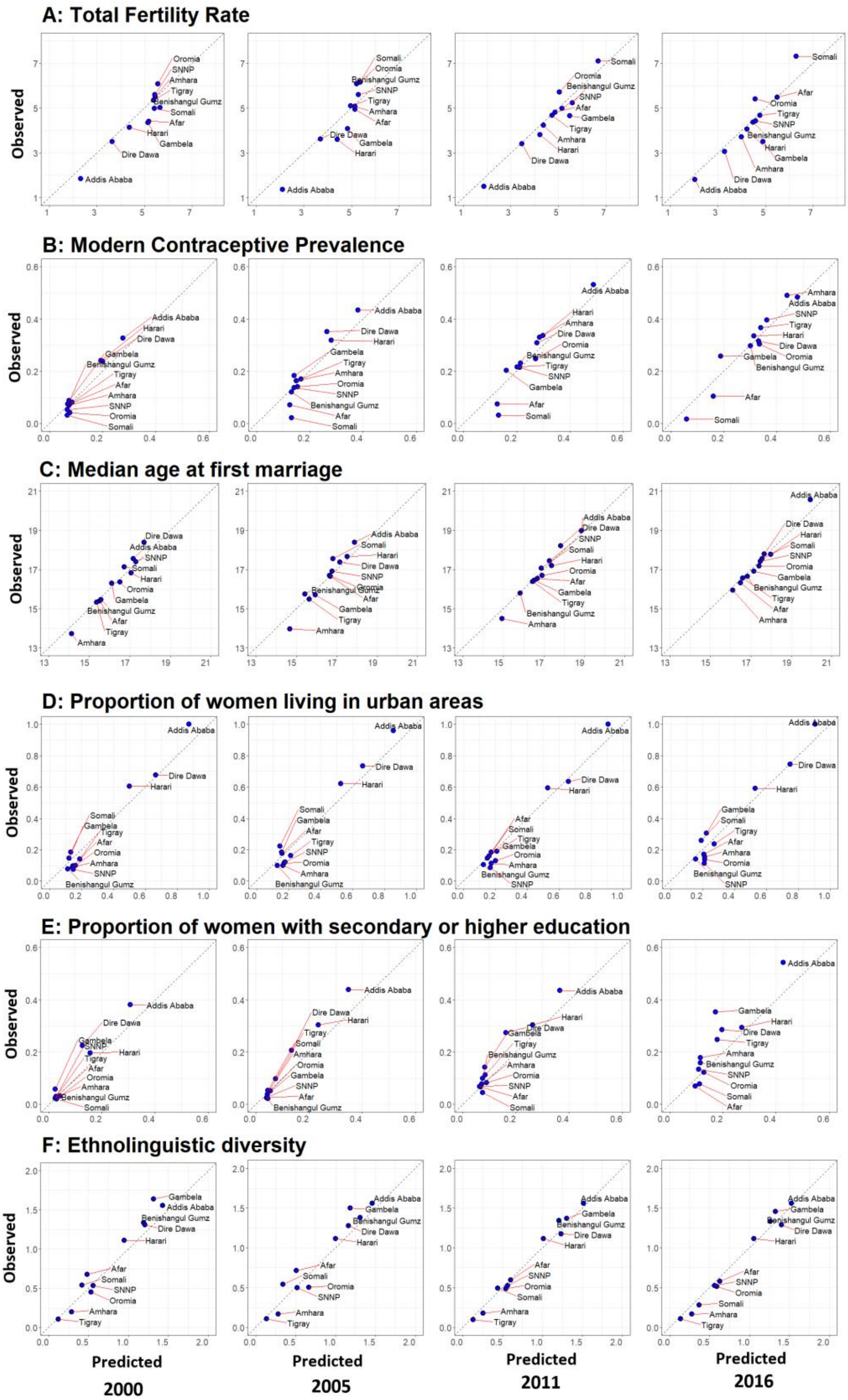


Figure 4.2. Observed and predicted regional values of the study variables for 2000-2016

A: Total fertility rates; **B**. modern Contraceptive prevalence (mCP); **C**. Median age at first marriage; **D**. Proportion of women living in urban areas; **E**. Proportion of women with secondary or higher education; **F**. Ethnolinguistic diversity at the zonal level. Note: The dashed line represents the 1:1 line

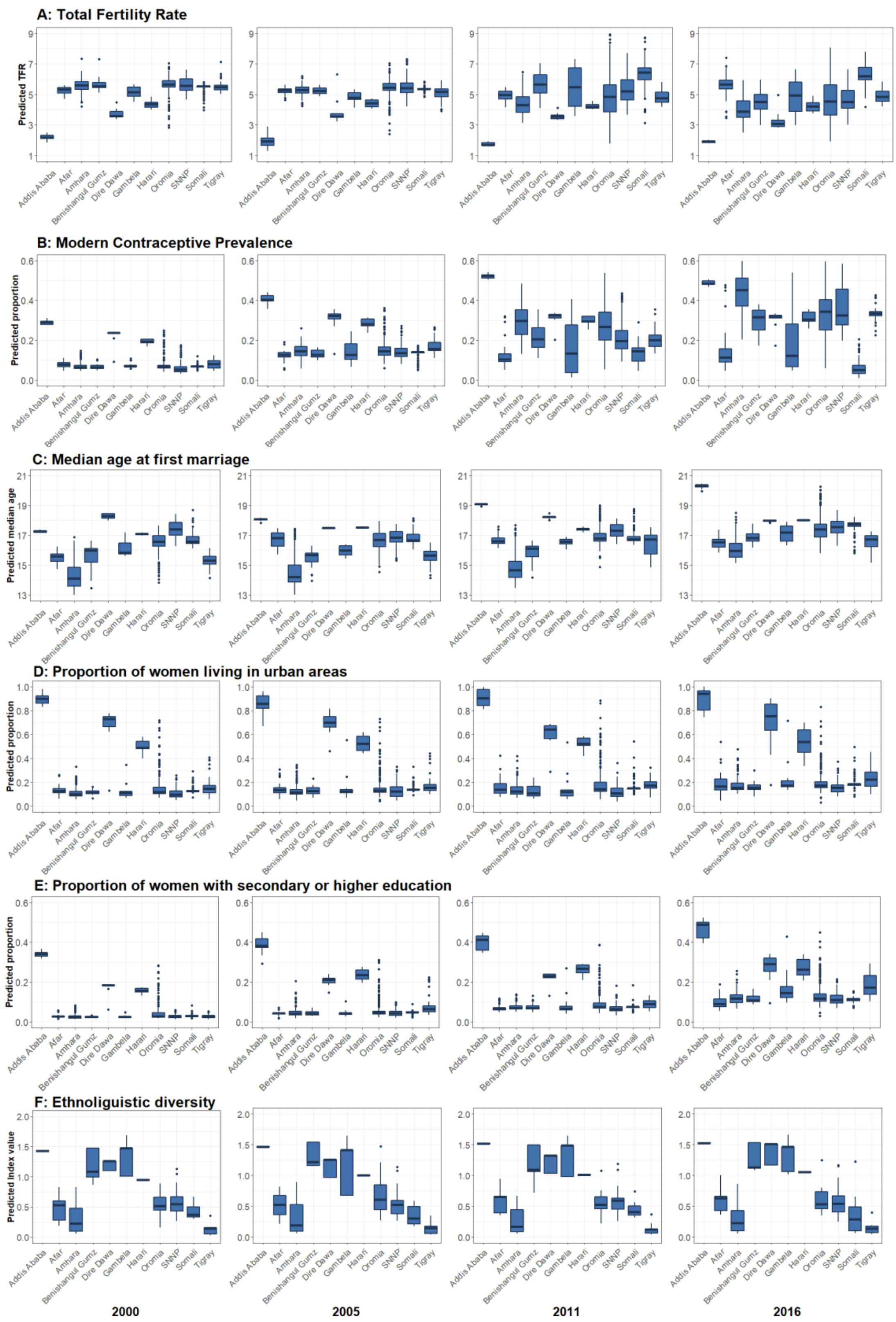


Figure 4.3. Predicted values of the study variables for each region

A. Total fertility rates; **B**. modern Contraceptive prevalence (mCP); C. Median age at first marriage; **D**. Proportion of women living in urban areas; **E**. Proportion of women with secondary or higher education; **F**. Ethnolinguistic diversity at the zonal level. Note that box plots were constructed with predicted values of study variables at districts within each regional state

4.6.3. Changes in spatial pattern of total fertility rates and proximate determinants

Spatial patterns of fertility have become more pronounced in recent years (Figure 4.4A). In 2000 and 2005, fertility levels were relatively uniform across Ethiopia, excluding the capital. However, by 2011, lower fertility levels started to emerge in the central and western parts of the country. In 2016, lower fertility extended further into the northern and western regions of Ethiopia. Likewise, spatial patterns of mCPs have become more pronounced in recent years (Figure 4.4B). While mCPs were generally low in all regional states other than the capital in 2000 and 2005, there has been a substantial increase in mCPs in many districts of the SNNP, Amhara, Oromia, and Benshangul-Gumuz regions since then. Additionally, the median age at first marriage has noticeably increased in the capital and the Amhara regional state between 2000 and 2016 (Figure 4.4C).

4.6.4. Changes in spatial pattern of distal determinants

Spatial patterning of the proportion of women living in urban areas remained low at the national level, excluding the multi-ethnic cities (Addis Ababa and Dire-Dawa), between 2000 and 2016 (Figure 4.4D). Similar spatial patterns were observed for the entire period in the proportions of women with secondary or higher education, where the spatial patterning remained low except in the multi-ethnic cities (Addis Ababa and Dire-Dawa) between 2000 and 2011. But the proportions of women with secondary or higher education increased in Tigray and Gambela regional states and western parts of Oromia regional states by 2016 (Figure 4.4E). Furthermore, spatial patterns of ethnolinguistic diversity at the zonal level revealed that zones tend to have a relatively high degree of ethnolinguistic heterogeneity in Benishangul-Gumz, Gambela regional states and the multi-ethnic cities (Addis Ababa and Dire-Dawa). On the other hand, zones tend toward ethnolinguistic homogeneity in Tigray, Amhara, and Somali regional states (Figure 4.4F).

4.6.5. Model-based Uncertainty

Our result showed that 95% credible intervals of uncertainty were often higher in Somali and Afar regional states and southern parts of Ethiopia (Figure 4.5). MBG credible intervals of uncertainty grow where there is heterogeneity in data, such as wide variation in TFR, or when there are no data to support prediction (Mayala et al., 2020) Hence, despite the smaller number of PSUs in 2000 and 2005 EDHS than 2011 and 2016 EDHS, lower credible intervals of TFR in 2000 and 2005 were observed, probably, due to limited heterogeneity in TFR. On the other hand, despite the larger number of PSUs in 2011 and 2016 EDHS than 2000 and 2005 EDHS, higher credible intervals of TFR in 2011 and 2016 were observed due to greater heterogeneity in TFR (Figure. 4.5A). In contrast, higher credible interval of median age at first marriage was observed in 2000 due to both greater heterogeneity in median age at first marriage and the lack of PSUs in Somali regional state. On the other hand, lower credible interval of median age at first marriage was observed in 2016 due to both less heterogeneity in median age at first marriage and the increased number of PSUs in Somali regional state (Figure. 4.5C).

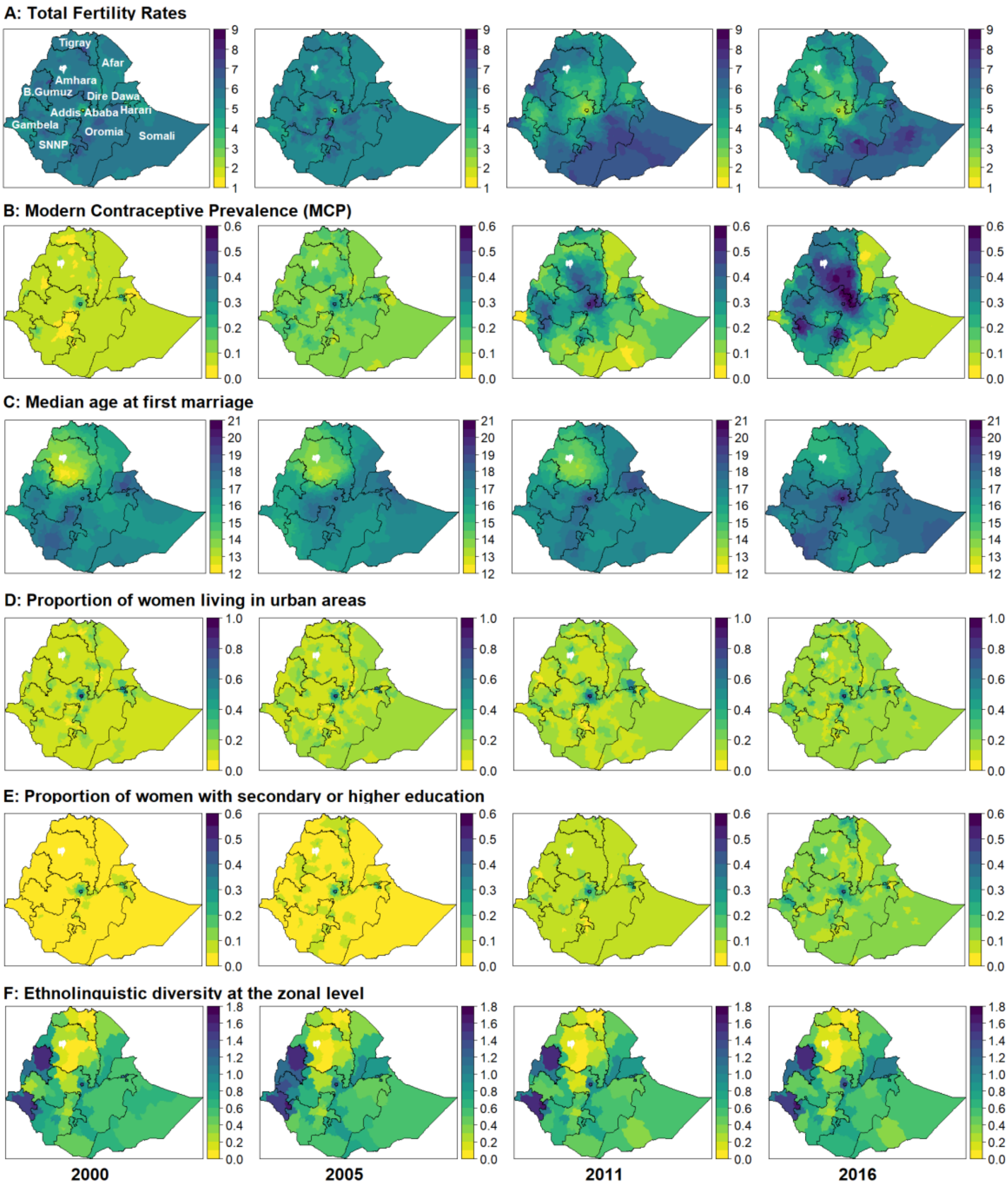


Figure 4.4. District-level changes in spatial and temporal patterns
A. total fertility rates (TFRs), **B**. modern contraceptive prevalence (mCP), **C**. median age at first marriage, **D**. proportion of women living in urban areas; **E**. proportion of women with secondary or higher education; **F**. ethnolinguistic diversity at the zonal level for 2000 – 2016 periods

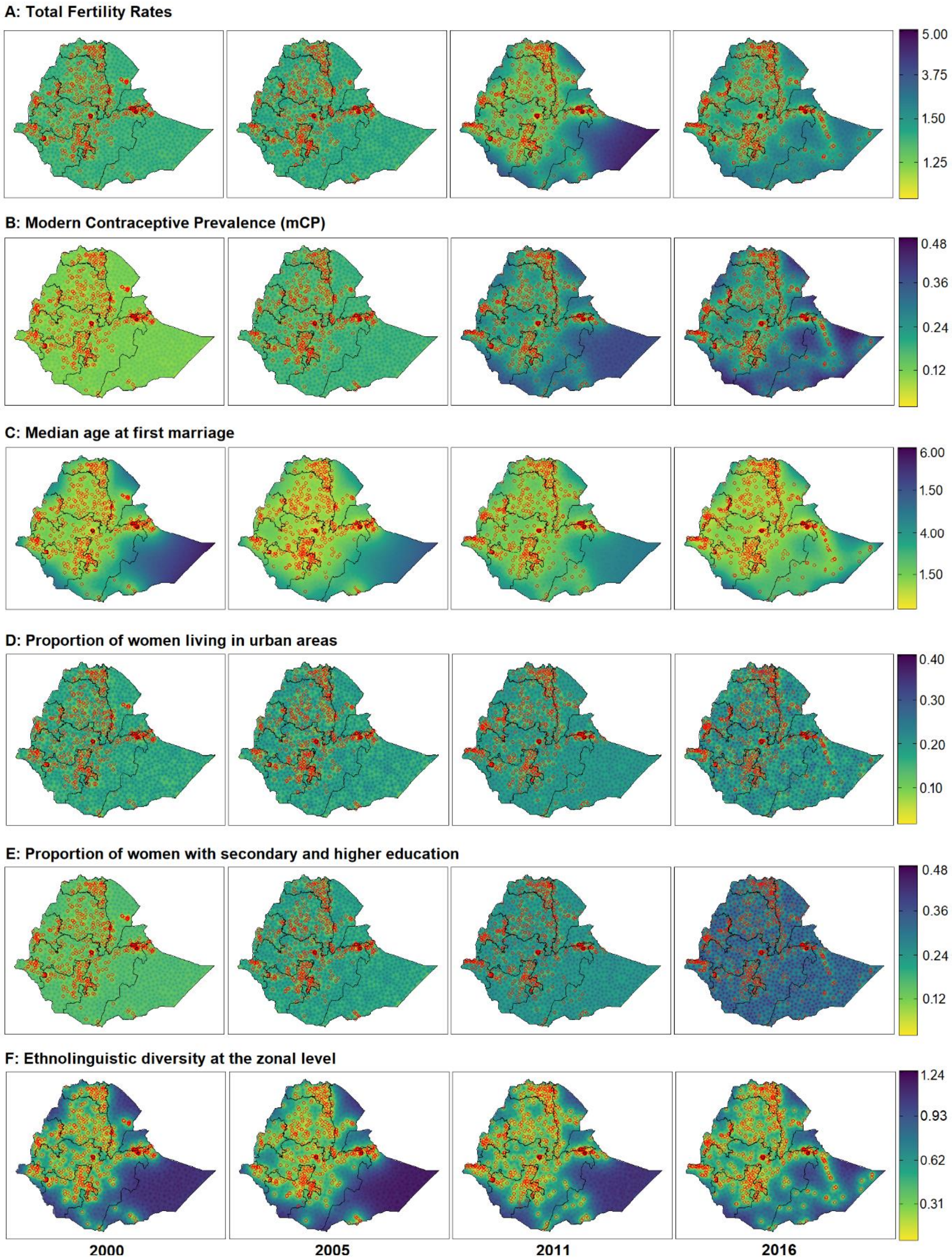


Figure 4.5. Model-based uncertainty measured by the width of the 95% credible interval
A. total fertility rates (TFRs), **B**. modern contraceptive prevanlence (mCPs), **C**. median age at first marriage, **D**. proportion of women living in urban areas; **E**.proportion of women with secondary or higher education; **F**.ethnolinguistic diversity at the zonal level for 2000 – 2016 periods. Red points are referred to locations of Primary Sampling Units (PSUs)

4.7. Discussion

This study shows the growing importance in Ethiopia of exploring the geographical distribution of fertility and its key determinants at the district level. Conducting a geographically disaggregated analysis can uncover spatial patterns in fertility that may be overlooked in aggregate-level analyses. In particular, this study elucidates three crucial spatial aspects of fertility in Ethiopia: a) the emergence of geographical variations in fertility between and within regional states, b) the development of distinct spatial patterns in the Total Fertility Rate (TFR) that closely mirror changes in spatial patterns of the modern contraceptive prevalence (mCP), and c) the combined influence of adaptation and diffusion effects on geographic variations in fertility at the district level.

First, our district-level analysis clearly shows that geographical variations in fertility have widened since 2011, both between regional states and, even more significantly, within regional states in Ethiopia. This highlights the emerging differentiation of fertility and its key determinants across the country. Previous studies have already indicated substantial geographical differentials in fertility between rural and urban areas, as well as among regional states in Ethiopia (Tessema et al., 2020, Desta, 2019). Our study further confirms the importance of sub-regional variations in fertility as a critical spatial aspect in recent years. We observed the emergence of geographical variations in district-level fertility within regions, particularly evident in the 2011 and 2016 Ethiopian Demographic and Health Surveys (EDHS). These findings imply that relying solely on national or regional data is insufficient to capture the contemporary geographical variations in fertility in Ethiopia. Previous research has highlighted that understanding and addressing such inter-district differentials in fertility is crucial for effective local-level planning, specifically in the provision of family planning, education, and healthcare services (Evans and Gray, 2018, Haque et al., 2019). Under Ethiopia’s decentralised health care system, District health offices are in charge of carrying out national health policies such the provision of equitable and high-quality healthcare through Ethiopia's decentralised healthcare system. Although the decentralisation of the health sector in Ethiopia aims to increase in equity in healthcare by improving responsiveness to needs in districts, it may deteriorate disparities between districts and within regional states in the absence of sufficient local data, clear guidelines, and effective monitoring. (Bergen et al., 2019). Therefore, our findings have important policy implications in terms of providing evidence for woreda based health sector planning in pursuit of narrowing health disparities between districts.

Secondly, our analysis reveals that changes in spatial patterns of district-level fertility in Ethiopia between 2000 and 2016 were closely aligned with changes in the proximate determinants, particularly the modern contraceptive prevalence (mCP), rather than the distal determinants. These findings indicate clear shifts in the spatial distribution of fertility across the country during this period.

In 2000 and 2005, lower fertility levels were primarily concentrated in the central part of Ethiopia. However, in 2011 and 2016, we observed a gradual spread of lower fertility from the central regions to the northern and western parts of the country. In contrast, higher fertility levels persisted consistently in the eastern parts, specifically in Afar and Somali regional states, throughout the entire 2000-2016 period. While higher proportions of women living in urban areas and with secondary education were consistently observed around multi-ethnic cities, where total fertility rates (TFRs) were significantly lower compared to other regions, the spatial patterns in district-level fertility corresponded more closely to the spatial patterns of mCP. Furthermore, these similar spatial patterns of TFR and mCP were independent of the spatial patterns of the selected socioeconomic determinants. This suggests that recent geographical variations in fertility at the district level in Ethiopia are broadly associated with variations in modern contraceptive use across the country. It also implies that mCP is adopted by women with diverse educational and residential backgrounds, including those who are less educated or live in rural areas. Previous studies have highlighted Ethiopia's successful implementation of a family planning (FP) programme through the Health Extension Programme (HEP), which delivers essential health services, including family planning, through health extension workers in marginalised and rural areas (Olson and Piller, 2013, Halperin, 2014). Our analysis shows substantial changes in the spatial patterns of TFR and mCP after 2005, coinciding with the implementation of the HEP since 2004 (Halperin, 2014, May and Rotenberg, 2020). It is worth noting that the performance of the HEP varied across regions, with better outcomes observed in the Amhara regional state (UNICEF, 2010), while challenges related to healthcare access, including the HEP, were encountered in regions such as Afar and Somali due to the nomadic lifestyle of pastoralist communities (Getnet et al., 2017). These regional disparities in the HEP are also reflected in our results, as Afar and Somali regional states exhibited persistently high TFR and low mCP levels from 2000 to 2016. This district-level analysis strongly supports the notion that FP programmes can facilitate fertility decline, and further demonstrates that geographic variations in the implementation of FP programmes can account for sub-national variations in contemporary fertility changes in sub-Saharan Africa.

Thirdly, our analysis reveals that lower fertility levels were consistently observed in highly urban areas such as Addis Ababa and Dire Dawa, suggesting that adaptation effects have played a role in encouraging lower fertility in these urban settings. However, the adaptation approach alone cannot fully explain the changes in spatial patterns of fertility in Ethiopia observed in 2011 and 2016. The diffusionist approach argues that cultural homogeneity, particularly through shared languages, can facilitate the diffusion of attitudes and information that support modern reproductive ideas and behaviours (Cleland and Wilson, 1987, Watkins, 1987). This phenomenon was notably observed in many districts of the Amhara regional state. As discussed earlier, fertility declines in districts within the Amhara regional state may have been facilitated by an effective family planning program, even in the presence of low proportions of women living in urban areas and with secondary education. In addition to this, the linguistic aspect may have also contributed to the spatial diffusion of behavioural change, such as the uptake of modern contraceptive methods and delays in first marriage, which could have facilitated fertility decline in the Amhara region. The Amhara regional state, accounting for approximately 23% of the total population in Ethiopia, is ethnolinguistically homogeneous, with around 92% of the population belonging to the Amhara ethnolinguistic group. Until 2020, Amharic was the sole working language of the federal government and the primary language of education in Ethiopia, among the more than 90 languages spoken in the country. Therefore some argue that the dominance of Amharic as the working language may undermine equity in Ethiopian society and create barriers to accessing health information in non-Amharic regions (Fufa Dugassa, 2006, Smith, 2008). Paradoxically, this suggests that the Amharic-speaking population in the Amhara regional state may have enjoyed the advantage of learning in their own language, which could facilitate public health communication and contribute to significant progress in increasing modern contraceptive uptake and median age at first marriage. Furthermore, the spatial spread of lower fertility rates and higher modern contraceptive prevalence around Addis Ababa, in particular, supports the view that cultural diversification in urban areas can accelerate the diffusion of fertility changes. Communication networks in urban areas often transcend socioeconomic and cultural boundaries, creating an environment conducive to innovative reproductive behaviours (Kulu, 2005, Goldstein, 1973, Lee and Farber, 1984, Lerch, 2019, Bongaarts and Watkins, 1996, Klüsener et al., 2019). Overall, our results support the view that adaptation and diffusion effects jointly influence the geographical variation in fertility at the district level in contemporary society, and this is confirmed in the context of Ethiopia as well(Salvati et al., 2020, Vitali and Billari, 2017).

This study has several limitations that should be acknowledged. Firstly, the modelled estimates are associated with a certain degree of uncertainty, which is particularly pronounced in the Somali and Afar regional states. This uncertainty may be attributed to smaller sample sizes and the exclusion of PSU without GPS in these two regions. Therefore, caution should be exercised when interpreting the results from these areas. Secondly, the aggregation method used in this study relies on simple mean statistics to aggregate the point estimate model surface to districts. While previous studies have utilized population-weighted mean statistics using datasets such as the Worldpop for aggregating up to administrative areas, this study used simple mean statistics due to the lack of population data specifically for women of childbearing age (15-49 years old) between 2000 and 2016. As Ethiopia's districts are relatively small in land area, we assumed a constant population density of women aged 15-49 within districts, which may introduce some limitations in accurately estimating fertility rates at the district level. In addition, as mentioned earlier, we applied survey weights to measure each study variable, following the guidelines outlined in the DHS manuals, in order to account for the complex survey design of the EDHS. This helps to ensure the representativeness of the findings at the national and sub-national levels.

Despite these limitations, our study has aimed to shed light on the changes in spatial patterns of fertility and the key proximate and distal determinants at the district level in Ethiopia between 2000 and 2016. While our focus in this study was primarily on describing the geographical distribution of fertility, the next chapter will delve deeper into the analysis by exploring the spatial correlation and heterogeneity of district-level Total Fertility Rates (TFRs) in relation to both proximate and distal determinants. This will be achieved through the use of spatial models, allowing us to examine the interplay between various factors and their influence on fertility outcomes across different districts.

4.8. Conclusion

The use of Bayesian MBG can offer a more nuanced picture of geographical variations in fertility and key determinants at the district level in Ethiopia. Our study shows that geographical variations in fertility at the district level have noticeably widened in recent years in Ethiopia. This result implies that focusing solely on national or regional data provides an inadequate description of geographical variations in fertility of contemporary Ethiopia. On the one hand, this is likely to impede the planning and monitoring of government programmes, while on the other, it is likely to obstruct efforts to explain the factors determining fertility in the country. The spatial pattens of fertility in recent years are related with geographical variations in socioeconomic and ethnolinguistic conditions but, primarily, with the use of modern contraception. The findings of this study have important policy implications, as they enable the visualization of demographic and health indicator changes across 981 districts. This information enables more informed decision-making for initiatives such as the Woreda-Based Health Sector Planning (WBHSP) and the Health Extension Program (HEP) in Ethiopia

4.9. References

ABBINK, J. 2011. Ethnic-based federalism and ethnicity in Ethiopia: reassessing the experiment after 20 years. *Journal of Eastern African Studies,* 5**,** 596-618.

ALEMAYEHU, M., LEMMA, H., ABRHA, K., ADAMA, Y., FISSEHA, G., YEBYO, H., GEBEYE, E., NEGASH, K., YOUSUF, J. & FANTU, T. 2016. Family planning use and associated factors among pastoralist community of afar region, eastern Ethiopia. *BMC women's health,* 16**,** 1-9.

BERGEN, N., RUCKERT, A., KULKARNI, M. A., ABEBE, L., MORANKAR, S. & LABONTë, R. 2019. Subnational health management and the advancement of health equity: a case study of Ethiopia. *Global health research and policy,* 4**,** 1-13.

BHATTACHARJEE, N. V., SCHAEFFER, L. E., MARCZAK, L. B., ROSS, J. M., SWARTZ, S. J., ALBRIGHT, J., GARDNER, W. M., SHIELDS, C., SLIGAR, A. & SCHIPP, M. F. 2019. Mapping exclusive breastfeeding in Africa between 2000 and 2017. *Nature medicine,* 25**,** 1205-1212.

BONGAARTS, J. 2015. Modeling the fertility impact of the proximate determinants: Time for a tune-up. *Demographic Research,* 33**,** 535-560.

BONGAARTS, J. & WATKINS, S. C. 1996. Social interactions and contemporary fertility transitions. *Population and development review***,** 639-682.

BURGERT-BRUCKER, C. R., DONTAMSETTI, T. & GETHING, P. W. 2018. The DHS program's modeled surfaces spatial datasets. *Studies in family planning,* 49**,** 87-92.

BURKE, M., HEFT-NEAL, S. & BENDAVID, E. 2016. Sources of variation in under-5 mortality across sub-Saharan Africa: a spatial analysis. *The Lancet Global Health,* 4**,** e936-e945.

CAMPISI, N., KULU, H., MIKOLAI, J., KLþSENER, S. & MYRSKYLæ, M. 2020. Spatial variation in fertility across Europe: Patterns and determinants. *Population, Space and Place***,** e2308.

CHEKOLE, M. K., KAHSAY, Z. H., MEDHANYIE, A. A., GEBRESLASSIE, M. A. & BEZABH, A. M. 2019. Husbands' involvement in family planning use and its associated factors in pastoralist communities of Afar, Ethiopia. *Reproductive health,* 16**,** 1-7.

CLELAND, J. & WILSON, C. 1987. Demand theories of the fertility transition: An iconoclastic view. *Population studies,* 41**,** 5-30.

CROFT, T. N., MARSHALL, A. M., ALLEN, C. K., ARNOLD, F., ASSAF, S. & BALIAN, S. 2018. Guide to DHS statistics. *Rockville: ICF,* 645.

DE BROE, S. & HINDE, A. 2006. Diversity in fertility patterns in Guatemala. *Population, Space and Place,* 12**,** 435-459.

DESTA, C. G. 2019. Rural-urban differential in fertility and maternal work participation: A study in Northwestern Ethiopia. *African Population Studies,* 33.

DWYER-LINDGREN, L., CORK, M. A., SLIGAR, A., STEUBEN, K. M., WILSON, K. F., PROVOST, N. R., MAYALA, B. K., VANDERHEIDE, J. D., COLLISON, M. L. & HALL, J. B. 2019. Mapping HIV prevalence in sub-Saharan Africa between 2000 and 2017. *Nature,* 570**,** 189-193.

EVANS, A. & GRAY, E. 2018. Modelling variation in fertility rates using geographically weighted regression. *Spatial Demography,* 6**,** 121-140.

FISH, T. D., JANOCHA, B., DONTAMSETTI, T. & MAYALA, B. K. 2020. Geospatial covariates: proxies for mapping urban-related indicators. *DHS Spatial Analysis Reports No. 19.* Rockville, Maryland, USA: ICF.

FUFA DUGASSA, B. 2006. Ethiopian language policy and health promotion in Oromia. *J. Soc. & Soc. Welfare,* 33**,** 69.

GETNET, F., HASHI, A., MOHAMUD, S., MOWLID, H. & KLINKENBERG, E. 2017. Low contribution of health extension workers in identification of persons with presumptive pulmonary tuberculosis in Ethiopian Somali Region pastoralists. *BMC health services research,* 17**,** 1-9.

GIORGI, E., FRONTERRê, C., MACHARIA, P. M., ALEGANA, V. A., SNOW, R. W. & DIGGLE, P. J. 2021. Model building and assessment of the impact of covariates for disease prevalence mapping in low-resource settings: to explain and to predict. *Journal of the Royal Society Interface,* 18**,** 20210104.

GOLDSTEIN, S. 1973. Interrelations between migration and fertility in Thailand. *Demography,* 10**,** 225-241.

GRAETZ, N., FRIEDMAN, J., OSGOOD-ZIMMERMAN, A., BURSTEIN, R., BIEHL, M. H., SHIELDS, C., MOSSER, J. F., CASEY, D. C., DESHPANDE, A. & EARL, L. 2018. Mapping local variation in educational attainment across Africa. *Nature,* 555**,** 48.

GURMU, E. & ETANA, D. 2014. Age at first marriage and first birth interval in Ethiopia: analysis of the roles of social and demographic factors. *African Population Studies,* 28**,** 1332-1344.

HALPERIN, D. T. 2014. Scaling up of family planning in low-income countries: lessons from Ethiopia. *The Lancet,* 383**,** 1264-1267.

HAQUE, I., DAS, D. & PATEL, P. P. 2019. Reading the geography of India's district-level fertility differentials: a spatial econometric approach. *Journal of Biosocial Science,* 51**,** 745-774.

HOGAN, D. P. & BIRATU, B. 2004. Social identity and community effects on contraceptive use and intentions in southern Ethiopia. *Studies in Family Planning,* 35**,** 79-90.

HOSSEINPOOR, A. R., BERGEN, N., SCHLOTHEUBER, A. & GROVE, J. 2018. Measuring health inequalities in the context of sustainable development goals. *Bulletin of the World Health Organization,* 96**,** 654.

HUANG, J., MALONE, B. P., MINASNY, B., MCBRATNEY, A. B. & TRIANTAFILIS, J. 2017. Evaluating a Bayesian modelling approach (INLA-SPDE) for environmental mapping. *Science of the Total Environment,* 609**,** 621-632.

ICF International and Central Statistical Agency. Ethiopia DHS 2000. Addis Ababa, Ethiopia and Calverton, MD: Central Statistical Agency, ICF International, 2000

ICF International and Central Statistical Agency. Ethiopia DHS 2005. Addis Ababa, Ethiopia and Calverton, MD: Central Statistical Agency, ICF International, 2005.

ICF International and Kenya National Bureau of Statistics. Kenya DHS 2014. Nairobi, Kenya and Calverton, MD: Kenya National Bureau of Statistics, ICF International, 2014.

ICF International and Central Statistical Agency. Ethiopia DHS 2016. Addis Ababa, Ethiopia and Calverton, MD: Central Statistical Agency, ICF International, 2016.

JANOCHA, B., DONOHUE, R. E., FISH, T. D., MAYALA, B. K. & CROFT, T. N. 2021. Guidance and recommendations for the use of indicator estimates at subnational administrative level 2. *DHS Spatial Analysis Report 20.* Rockville, Maryland, USA: ICF.

KLþSENER, S., DRIBE, M. & SCALONE, F. 2019. Spatial and social distance at the onset of the fertility transition: Sweden, 1880–1900. *Demography,* 56**,** 169-199.

KRAINSKI, E. T., GõMEZ-RUBIO, V., BAKKA, H., LENZI, A., CASTRO-CAMILO, D., SIMPSON, D., LINDGREN, F. & RUE, H. 2018. *Advanced spatial modeling with stochastic partial differential equations using R and INLA*, CRC Press.

KULU, H. 2005. Migration and fertility: Competing hypotheses re-examined. *European Journal of Population/Revue européenne de Démographie,* 21**,** 51-87.

LAELAGO, T., HABTU, Y. & YOHANNES, S. 2019. Proximate determinants of fertility in Ethiopia; an application of revised Bongaarts model. *Reproductive health,* 16**,** 13.

LEE, B. S. & FARBER, S. C. 1984. Fertility adaptation by rural—urban migrants in developing countries: The case of Korea. *Population studies,* 38**,** 141-155.

LERCH, M. 2019. Regional variations in the rural-urban fertility gradient in the global South. *PloS one,* 14**,** e0219624.

LI, Z., HSIAO, Y., GODWIN, J., MARTIN, B. D., WAKEFIELD, J., CLARK, S. J., ESTIMATION, W. S. F. T. U. N. I.-A. G. F. C. M. & GROUP, I. T. A. 2019b. Changes in the spatial distribution of the under-five mortality rate: Small-area analysis of 122 DHS surveys in 262 subregions of 35 countries in Africa. *PloS one,* 14**,** e0210645.

LINDGREN, F. & RUE, H. 2015. Bayesian spatial modelling with R-INLA. *Journal of Statistical Software,* 63**,** 1-25.

MAY, J. F. & ROTENBERG, S. 2020. A Call for Better Integrated Policies to Accelerate the Fertility Decline in Sub-Saharan Africa. *Studies in Family Planning,* 51**,** 193-204.

MAYALA, B. K., BHATT, S. & GETHING, P. 2020. Predicting HIV/AIDS at subnational levels using DHS covariates related to HIV. *DHS Spatial Analysis Reports No. 18.* Rockville, Maryland, USA: ICF.

MAYALA, B. K., DONTAMSETTI, T., FISH, T. D. & CROFT, T. 2019a. *Interpolation of DHS survey data at subnational administrative Level 2,* DHS Program, ICF.

MAYALA, B. K., DONTAMSETTI, T., FISH, T. D. & CROFT, T. N. 2019b. Interpolation of DHS survey data at subnational administrative Level 2. *DHS Spatial Analysis Reports No. 17.* Rockville, Maryland, USA: ICF.

MOH 2015. Health sector transformation plan (2015/16–2019/20). Federal Ministry of Health Addis Ababa, Ethiopia.

MOH 2021. Health Sector Transformation Plan (2020/21-2024/25). Federal Ministry of Health Addis Ababa, Ethiopia.

OLSON, D. J. & PILLER, A. 2013. Ethiopia: an emerging family planning success story. *Studies in Family Planning,* 44**,** 445-459.

OSGOOD-ZIMMERMAN, A., MILLEAR, A. I., STUBBS, R. W., SHIELDS, C., PICKERING, B. V., EARL, L., GRAETZ, N., KINYOKI, D. K., RAY, S. E. & BHATT, S. 2018. Mapping child growth failure in Africa between 2000 and 2015. *Nature,* 555**,** 41-47.

PEZZULO, C., NILSEN, K., CARIOLI, A., TEJEDOR-GARAVITO, N., HANSPAL, S. E., HILBER, T., JAMES, W. H., RUKTANONCHAI, C. W., ALEGANA, V. & SORICHETTA, A. 2021. Geographical distribution of fertility rates in 70 low-income, lower-middle-income, and upper-middle-income countries, 2010–16: a subnational analysis of cross-sectional surveys. *The Lancet Global Health,* 9**,** e802-e812.

REINER JR, R. C., WIENS, K. E., DESHPANDE, A., BAUMANN, M. M., LINDSTEDT, P. A., BLACKER, B. F., TROEGER, C. E., EARL, L., MUNRO, S. B. & ABATE, D. 2020. Mapping geographical inequalities in childhood diarrhoeal morbidity and mortality in low-income and middle-income countries, 2000–17: analysis for the Global Burden of Disease Study 2017. *the lancet,* 395**,** 1779-1801.

SABATER, A. & GRAHAM, E. 2019. International migration and fertility variation in Spain during the economic recession: A spatial Durbin approach. *Applied Spatial Analysis and Policy,* 12**,** 515-546.

SALVATI, L., BENASSI, F., MICCOLI, S., RABIEI-DASTJERDI, H. & MATTHEWS, S. A. 2020. Spatial variability of total fertility rate and crude birth rate in a low-fertility country: patterns and trends in regional and local scale heterogeneity across Italy, 2002–2018. *Applied Geography,* 124**,** 102321.

SMITH, L. 2008. The politics of contemporary language policy in Ethiopia. *Journal of Developing Societies,* 24**,** 207-243.

TESSEMA, Z. T., AZANAW, M. M., BUKAYAW, Y. A. & GELAYE, K. A. 2020. Geographical variation in determinants of high-risk fertility behavior among reproductive age women in Ethiopia using the 2016 demographic and health survey: a geographically weighted regression analysis. *Archives of Public Health,* 78**,** 1-12.

UN 2015. Transforming our world: the 2030 agenda for sustainable development. UN General Assembly New York.

UNICEF 2010. *HEWs and Health Post Performance and Community Satisfaction and Perception: HEP Evaluation in Rural Ethiopia,* Addis Ababa, UNICEF.

UNOCHA. 2020. *Ethiopia: Administrative map* [Online]. UNOCHA. Available: https://reliefweb.int/map/ethiopia/ethiopia-administrative-map-october-2020 [Accessed October 2020].

VITALI, A. & BILLARI, F. C. 2017. Changing determinants of low fertility and diffusion: A spatial analysis for Italy. *Population, Space and Place,* 23**,** e1998.

WANG, D. & CHI, G. 2017. Different Places, Different Stories: A Study of Spatial Heterogeneity of County-Level Fertility in China. *Demographic research,* 37**,** 493.

WATKINS, S. C. The fertility transition: Europe and the Third World compared. Sociological forum, 1987. Springer, 645-673.
WATKINS, S. C. 1990. From local to national communities: The transformation of demographic regimes in Western Europe, 1870-1960. *Population and development review*, 241-272.
WHITE, M. J. 1986. Segregation and diversity measures in population distribution. *Population index*, 198-221.
ZON, H., PAVLOVA, M., DRABO, K. M. & GROOT, W. 2017. Municipal health services provision by local governments: a systematic review of experiences in decentralized Sub-Saharan African countries. *Health Policy and Planning*, 32**,** 1327-1336.

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1704116	**Title**	Mr
First Name(s)	Myunggu		
Surname/Family Name	Jung		
Thesis Title	Spatial aspects of fertility change in Ethiopia between 2000 and 2016: a district-level analysis		
Primary Supervisor	Kazuyo Machiyama		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Demographic Research
Please list the paper's authors in the intended authorship order:	Myunggu Jung, Christopher I Jarvis, Ian M Timæus, Kazuyo Machiyama
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I conceived of study design and the statistical analysis plan, conducted the statistical analysis and wrote the first draft with feedback from co-authors.

SECTION E

Student Signature	[redacted]
Date	20/12/2022

Supervisor Signature	[redacted]
Date	20/12/2022

Chapter 5

Paper 2: Spatial dependence and heterogeneity of fertility in Ethiopia between 2000 and 2016: a district level analysis

Paper 2: Spatial dependence and heterogeneity of fertility in Ethiopia between 2000 and 2016: a district level analysis

Myunggu Jung[1], Christopher I Jarvis[2], Ian M Timæus[1,3], Kazuyo Machiyama[1]

1. Department of Population Health, London School of Hygiene & Tropical Medicine, London, UK
2. Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK
3. Centre for Actuarial Research, University of Cape Town, Cape Town, South Africa

5. Chapter 5: Spatial dependence and heterogeneity of fertility in Ethiopia between 2000 and 2016: a district level analysis

5.1. Overview

In previous chapters, it has been highlighted that fertility studies in sub-Saharan African (SSA) countries, including Ethiopia, often overlook the spatial dimension when examining sub-national variations in fertility, despite theoretical and empirical evidence supporting the presence of spatial dependency and heterogeneity of fertility in most populations. Additionally, Chapter 4 demonstrated the existence of geographical variations within Ethiopia at the district-level in terms of total fertility rate (TFR) and selected proximate and distal determinants. Building upon this, Chapter 5 aims to compare non-spatial and spatial regression models to explore the spatial autocorrelation of district-level TFR and the spatial heterogeneity in the effects of both proximate and distal determinants of TFR in Ethiopia between 2000 and 2016. Furthermore, this chapter delves into the discussion of how the spatial location of districts, distances between them, and varying levels of proximate and distal determinants collectively contribute to the geographical variations observed in district-level TFRs in Ethiopia.

Objective 3	To assess effects of key selected proximate and distal determinants on geographical variations in fertility at the district level between 2000-2016 with a non-spatial model.
Objective 4	To assess spatial autocorrelation of district-level fertility by using a spatial model.
Objective 5	To explore spatial heterogeneity in relationships between TFRs and both proximate and distal determinants in Ethiopia by using geographically weighted regression between 2000 and 2016.

5.2. Role of candidate

I conceived of study design and the statistical analysis plan which was agreed by the co-authors. I conducted the statistical analysis and wrote the first draft of the manuscript with feedback and inputs provided from Christopher I Jarvis, Ian M Timæus and Kazuyo Machiyama.

5.3. Abstract

Background

There is growing evidence of spatial autocorrelation and heterogeneity in fertility patterns, influenced by various factors, at the district level in high- and middle-income countries. However, spatial modelling of geographical variations in district-level fertility in sub-Saharan Africa remains limited. From the perspective of demographic theories, geographical variation in fertility is viewed as a response to socioeconomic conditions (the adaptationist approach) or the spatial diffusion of social acceptance of fertility control (the diffusionist approach). This paper aims to explore spatial autocorrelation and heterogeneity in district-level fertility in relation to both proximate and distal determinants in Ethiopia between 2000 and 2016.

Methods

We used data from the 2000, 2005, 2011, and 2016 Ethiopia Demographic and Health Surveys to examine total fertility rates (TFRs) as well as two proximate determinants (modern contraceptive prevalence (mCP) and median age at first marriage) and three distal determinants (women living in urban areas and with secondary and higher education, ethnolinguistic diversity). We compared a spatial explicit regression model, the spatial lag model (SLM), with a non-spatial model to investigate spatial autocorrelation of TFRs. Additionally, we employed a geographically weighted regression (GWR) model to explore the spatially heterogeneous relationship between TFRs and the selected determinants at the district level in Ethiopia.

Results

Our results shows that the district-level data in Ethiopia were more consistent with the SLM compared to the non-spatial model in 2011 and 2016, suggesting that spatial autocorrelation of district-level TFR became stronger in recent years. The GWR model shows the relationships between fertility and proximate and distal determinants of fertility are spatially heterogeneous at the district level. Urban-rural fertility differences in Ethiopia were more closely associated with diverse socioeconomic conditions, while the spatial spread of lower fertility from Addis Ababa to the Amhara region was influenced by spatially heterogeneous effects of mCP, age at marriage, and ethnolinguistic diversity.

Conclusion

Space and place increasingly matter to recent geographical variations in fertility in Ethiopia. Socioeconomic and cultural characteristics, even in the same region, differ between districts and fertility in a district is affected by the spatial location of the district and the characteristics of nearby districts (diffusion effects) as well as by its own characteristics (adaptation effects). This study provides additional insights into how geographical location of and distance between districts as well as socioeconomic, cultural characteristics and reproductive behaviours in districts can jointly shape geographical variations in district-level fertility in Ethiopia.

5.4. Introduction

Among the sub-Saharan African (SSA) countries surveyed by the Demographic and Health Surveys (DHS) program, Ethiopia stands out with the largest fertility differences based on urban-rural residence and women's educational backgrounds. In 2016, the total fertility rate (TFR) in urban areas was 2.3, compared to 5.2 in rural areas. Similarly, women with a secondary education had a TFR of 2.1, while those with less than secondary education had a TFR of 5.0 (ICF, 2015). Ethiopia is also the second most populous country in sub-Saharan Africa, and it exhibits substantial variations in fertility across regions. In 2016, the highest TFR of 7.2 was observed in the Somali region, while the lowest TFR of 1.8 was recorded in Addis Ababa (ICF, 2016).

Demographers contend that comprehensive analysis of factors impacting fertility variation requires that a distinction be made between two determinants: proximate and distal determinants (Bongaarts, 1978). Distal determinants encompass socioeconomic and cultural factors that indirectly influence fertility only through their impact on the proximate determinants. On the other hand, proximate determinants refer to biological and behavioural factors that directly affect overall fertility rates (Davis and Blake, 1956, Bongaarts and Potter, 1983).

Ethiopia has witnessed the substantial socioeconomic and cultural differences across eleven regional states. The country comprises three urban regions (Addis Ababa, Dire-Dawa, and Harari) alongside eight other regions. The three urban regions are characterised by their multi-ethnic composition. Among the remaining eight regions, two western regions (Benishangul Gumuz and Gambela) and two eastern regions (Afar and Somali) are commonly classified as Developing Regional States (DRS). These DRS predominantly consist of pastoral and migratory communities, which often experience worse socioeconomic and health outcomes compared to the national average. Of the eight regions, five (Tigray, Amhara, Oromia, Afar, and Somali) are ethnolinguistically homogeneous regions. The names of these regions correspond to the names of the majority ethnolinguistic groups residing within them, such as Tigrayian, Amhara, Oromo, Afar, and Somali ethnolinguistic groups. On the other hand, the other three regions (Benishangul-Gumuz, Gambela, and SNNP) are characterized as multi-ethnic regions where multiple ethnolinguistic groups coexist within the same geographic area.

Furthermore, there are significant regional disparities in the proximate determinants of fertility. Variation in the use of modern contraceptives exists among different regions in the country. For instance, the Somali regional state had the lowest modern contraceptive prevalence (1.4%) compared to the Amhara regional state (46.9%) in 2016. Additionally, the median age at first marriage plays a crucial role in regional fertility variations in Ethiopia, as childbearing outside marriage is uncommon (Gurmu and Etana, 2014, Reda and Lindstrom, 2014). Child marriage, defined as women entering marriage before the age of 18, is prevalent in Ethiopia and varies across regional states. The prevalence of early marriage before the age of 18 for women in Ethiopia (40%) exceeds the average in Eastern and Southern Africa (35%) and is approximately double the global average (21%) (UNICEF, 2018). Notably, the Amhara regional state has the highest prevalence of child marriage, with 74% of women married before the age of 18, significantly surpassing the national average of 40% in 2014 (Erulkar, 2013, Mekonnen et al., 2018) .

Previous studies have shown that substantial variations in fertility between populations and regions in Ethiopia, attributing them to differences in both proximate and distal determinants of fertility (Shiferaw et al., 2015, Tessema and Tamirat, 2020, Laelago et al., 2019). However, in addition to examining regional variations in fertility, acknowledging geographical variations in fertility at the district level is also important for Ethiopia's health policy planning and implementation. Districts, referred to as woredas in the Amharic language, represent the third level of administrative division in the country. The Woreda-Based Health Sector Annual Plan (WBHSP) plays a central role in Ethiopia's Health Sector Transformation Plan (HSTP) and guides the decentralised health planning at the woreda level (MoH, 2021). Each woreda annually prepares a woreda-based health plan through the WBHSP based on the relevant national priorities established by the federal Ministry of Health. The WBHSP is, then, combined to form an annual national health plan. The Ethiopian government at all levels and all other health partners adhere to this operational planning (Teshome and Hoebink, 2018).

Compared to previous studies conducted at the national or regional level, there has been limited research on spatial geographical variations in fertility at the district level in Ethiopia, mainly due to the scarcity of district-level data. However, recent studies conducted in high- and middle-income countries (These studies also highlight substantial heterogeneities in the association between local-level fertility and socioeconomic and cultural factors, both in terms of magnitude and direction, observed in various parts of the world) provide growing evidence that district-level fertility patterns are often influenced by spatial autocorrelation of fertility and spatially heterogeneous relationships with socioeconomic factors. In other words, a decline in fertility in a particular location is often associated with a decline in fertility in neighbouring areas (Salvati et al., 2020, Sabater and Graham, 2019, Singh et al., 2017). These studies also highlight substantial heterogeneities in the association between local-level fertility and socioeconomic and cultural factors, both in terms of magnitude and direction (Wang and Chi, 2017, Campisi et al., 2020, Haque et al., 2019, Evans and Gray, 2018, Goldstein and Klüsener, 2014). Therefore, these studies argued that there is no single universal explanation for geographical variations in fertility in relation to socioeconomic and cultural factors. Instead, successful policymaking requires an understanding of local fertility patterns and the underlying factors that drive these patterns.

From the perspective of demographic theories, geographical variations in fertility are often explained by adaptationist or diffusionist approaches. The adaptationist approach suggests that fertility differences across geographic areas stem from adaptations to distinct socioeconomic conditions in each area. On the other hand, the diffusionist approach argues that such variations are a result of the spread of information and social acceptability of fertility control, influenced by cultural and geographical distances rather than solely by socioeconomic factors (Carlsson, 1966, Bongaarts and Watkins, 1996). Increasing evidence suggests that the geographical spread of ideas and norms across local areas may not align perfectly with geographical variations in socioeconomic characteristics (Vitali and Billari, 2017, Costa et al., 2021). For instance, cultural and normative changes, such as the desired number of children or modern contraceptive use, may occur at a slower pace compared to rapid socioeconomic transformations in some areas, or vice versa. Consequently, areas with similar socioeconomic characteristics may exhibit different fertility outcomes due to their distinct local contexts (Staveteig et al., 2018). Moreover, recent fertility studies in HMICs, including, Italy (Vitali and Billari, 2017), Spain (Sabater and Graham, 2019), Europe (Campisi et al., 2020), China (Wang and Chi, 2017) and India(Haque et al., 2019), demonstrated that the adaption and diffusion effects can jointly account for sub-national fertility variations. However, these studies often focused on the association between distal determinants of fertility, such as socioeconomic and cultural factors, while neglecting the role of proximate determinants in explaining geographical variations in local-level fertility. In many SSA countries, including Ethiopia, the prevalence of contraceptive use, for example, remains low and exhibits geographic variation within the country. Therefore, exploring both proximate and distal determinants of fertility can enhance our understanding of geographical variations in fertility in SSA.

In this study, we examine whether and how geographical variations in district-level fertility are associated with spatial autocorrelation and heterogeneity of district-level fertility in relation to key selected proximate and distal determinants of fertility using data from the 2000, 2005, 2011, and 2016 EDHS surveys. To the best of our knowledge, this is the first study to utilize spatial models in examining the relationship between district-level fertility and its proximate and distal determinants in sub-Saharan Africa (SSA), including Ethiopia.

5.5. Methods

5.5.1. Data

We used Ethiopia Demographic and Health Surveys (EDHS) conducted in 2000, 2005, 2011 and 2016 (ICF, 2011, ICF, 2016, ICF, 2005, ICF, 2000). The EDHS surveys are nationally representative surveys that offer estimates for a variety of demographic and health variables which are comparable over time and space. In order to facilitate spatial analysis and protect participant privacy, the EDHS offers global positioning system (GPS) coordinates for primary sampling units (PSUs) of aggregated household survey data. In urban areas, these coordinates are displaced up to 2 km, and up to 5 km in in rural areas.

The outcome of interest for this study is the TFR at the district level. TFR refers to the average number of live births a woman would have if she was subject to the current age-specific fertility rates throughout her reproductive years (aged 15–49 years). The study includes two proximate determinants, proportion of married women currently using any method of contraception (mCP) and median age at first marriage. Three distal determinants are also included: the proportion of women with secondary or higher education, the proportion of women living in urban areas, and Ethnolinguistic diversity. Table 5.1 provides variable names with their descriptions. These explanatory variables were chosen as previous studies have shown them to be associated with geographical variation in fertility in Ethiopia (ICF, 2015, Laelago et al., 2019, Hogan and Biratu, 2004). Notably, Ethiopia has the largest fertility differentials in terms of urban-rural residence and women's education backgrounds among SSA countries surveyed by the DHS programme. Moreover, Ethiopia is one of the few SSA countries that the government strongly supports for family planning through the Health Extension Programme (HEP) since 2004. This program specifically targets rural communities and aims to increase the adoption of modern contraceptive methods (Olson and Piller, 2013). Additionally, as mentioned earlier, Ethiopia exhibits geographical variations in ethnolinguistic contexts, and the age at first marriage plays a significant role in facilitating recent fertility decline, particularly considering the relatively low proportion of births occurring outside marriage compared to other sub-Saharan African countries (Rogers and Stephenson, 2018, Sibanda, 2003).

Table 5.1. Explanatory variables used in this study

Theme	Variable name	Description
Outcome variable	Total Fertility Rates	Total fertility rate for the three years preceding the survey for age group 15-49 expressed per woman
Proximate determinants	Married women currently using any method of modern contraception	Proportion of currently married or in union women currently using any method of modern contraception (modern contraceptive prevalence (mCP))
	Median age at first marriage	Median age at first marriage or union in years among women age 15-49
Distal determinants	Women with secondary or higher education	Proportion of women with secondary or higher education which is equivalent to 9 or longer years of schooling
	Women living in urban areas	Proportion of women living in urban areas
	Index of ethnolinguistic diversity at the zone level	The index of diversity is known as the 'entropy index'. While lower values of the index (close to zero) indicate similarity in ethnolinguistic composition in a zone, larger values of the index show ethnolinguistic diversity.

To investigate the relationship between district-level TFR and the key selected proximate and distal determinants at the district level, I used estimates of these variables across 981 districts, as obtained from Chapter 4. In Chapter 4, we employed a Bayesian model-based geostatistics (MBG) approach, specifically the Integrated Nested Laplace Approximation and Stochastic Partial Differential

Equation (INLA-SPDE) method to estimate the levels of study variables at the district level (Rue et al., 2009). More detailed information on the Bayesian modelling approach using INLA-SPDE can be found in Chapters 3 and 4, as well as in other sources (Blangiardo and Cameletti, 2015). In brief, the Bayesian MBG with INLA-SPDE approach has been widely employed in previous studies utilizing data from Demographic and Health Surveys (DHS) to create high resolution maps of health indicators (Mayala et al., 2019b, Fish et al., 2020). Similarly, I adopted a similar modelling framework to estimate the levels of the selected study variables at the district level in Ethiopia. This involved two main steps: first, estimating the spatial distribution of the study variables at a high resolution, and second, aggregating the estimates from the pixel level of the model surface to the district polygons.

In this Chapter, we employed three different models to investigate the relationships between the key proximate and distal determinants and the outcome variable, which is the district-level TFR.

5.5.3. Statistical analysis

5.5.3.1 Spatial autocorrelation of TFR at Primary Sampling Units

To analyse the spatial autocorrelation of TFR at the DHS PSU level between 2000 and 2016, we conducted semi-variogram analysis. The semi-variogram describes the extent to which nearby locations exhibit similar values by measuring the semi-variance. Then a variogram $\gamma(h)$ of TFR, Z, at location of PSU, s, defined as;

$$\hat{\gamma}(h) = \frac{1}{2n}\sum_{i=1}^{N}\{Z(s_i + h) - Z(s_i)\}^2$$

where N represents the number of pair points of PSUs separated by lag distance h used to estimate the value of $\hat{\gamma}(h)$. This collection of points is commonly referred to as the variogram cloud, illustrating the semivariances between all pairs of points. If there is spatial autocorrelation of TFR at the PSU level, the semi-variance is expected to be small at relative short distance and tends to increase with distance, indicating that observations closer in distance tend to be more similar than those farther apart. We employed an exponential model by assuming that the semi-variance of TFR at the PSU level increases exponentially as the distance between locations increases due to the spatial autocorrelation of TFR. Specifically, we present a) TFR at the PSU level and b) the variogram cloud, and c) the empirical and theoretical semi-variogram plots for TFR by using an exponential model under the assumption of spatial autocorrelation.

5.5.3.2. Spatial autocorrelation and heterogeneity of TFR at the district level

In order to examine spatial autocorrelation and heterogeneity of district-level TFRs, I employed both global and local regression models. Global regression models make the assumption that the relationships between the outcome variable and explanatory variables do not vary across space, which is known as spatial stationarity. However, this assumption may not hold true in reality. On the other hand, local regression models take into account the possibility that the relationships between variables may vary across different locations, which is referred to as spatial non-stationarity. This means that the relationship between the variables may be different in different areas or districts. By utilising both global and local regression models, I aimed to capture both the overall spatial patterns and any localized variations in the relationship between district-level TFRs and the selected explanatory variables.

1) Global model

For the global regression model, I compare non-spatial and spatial linear regression models to investigate the presence of spatial autocorrelation in district-level TFRs.

a) Non-spatial linear regression

In the non-spatial linear regression model, the relationships between the district-level TFR (outcome variable) and a set of explanatory variables are investigated. The model assumes that the observations at the district level are independent of each other and do not consider any potential spatial correlation. The model is represented by the equation:

$$y(TFR)_j = \beta_0 + \sum\nolimits_{m=1}^{N} \beta_m X_{mj} + \varepsilon_j$$

where j denotes the i^{th} districts, where i = 1 to 981; y_i is the outcome variable, TFR; β_m is the regression coefficient for explanatory variable . m denotes the selected explanatory variables, including mCP, median age at first marriage, proportion of urban population and women with secondary and higher education and ethnolinguistic homogeneity at Zonal areas. X represent the value of explanatory variable m at district j. ε_j is an error term for the regression equation. The non-spatial linear regression model generally optimises regression coefficients (β) by minimising the sum of squared prediction errors (Anselin and Arribas-Bel, 2013). In addition, the non-spatial linear regression model assumes that observations at the district level are independent of each other and does not account for potential spatial correlation. However, previous have demonstrated the presence of spatial autocorrelation in fertility (Vitali and Billari, 2017, Costa et al., 2021). These spatial interactions are not accounted for in non-spatial models, leading to a potential mis-specification of the non-spatial linear regression model (Anselin and Arribas-Bel, 2013).

b) Spatial linear regression - Spatial Lag model (SLM)

The spatial lag model (SLM) incorporates spatial autocorrelation into the regression model by including a 'spatially-lagged outcome variable.' It assumes a spatial relationship between the outcome variable and the explanatory variables (Ward and Gleditsch, 2018). The SLM assumes that the outcome variable in one district is influenced by the outcome variable of neighboring districts. The SLM is represented by the following equation:

$$y(TFR)_j = \beta_0 + \rho \sum\nolimits_{cn=1}^{N} W_j y_{cn} + \sum\nolimits_{m=1}^{N} \beta_m X_{mj} + \varepsilon_j$$

ρ is the spatial lag term captures the extent of spatial autocorrelation in fertility. A higher value of ρ indicates a stronger spatial autocorrelation (Kostov, 2010). cn refers to the connected neighbouring districts and y_{cn} is the TFR of the neighbouring district. X represent the value of explanatory variable m at location j. W_j denotes the spatial weight for a given district j. The weight matrix was generated based on first-order Queens' contiguity (Figure 5.2). In the sense of Queen contiguity, it assigns a binary spatial weight (0,1) to any connected neighbouring districts. This corresponds to the Queen's movement in chess. More detailed information about weight is described in Chapter 3 with Figure 3.2.

Figure 5.1. Queen's case contiguity of (A) first and (B) second order

5.5.3.3. Local Model - Geographically Weighted Regression (GWR)

Global regression models assume spatial stationarity, meaning that the relationships between the outcome variable and explanatory variables are assumed to be constant across all locations (Brunsdon et al., 1998, Brunsdon et al., 1996). In contrast, Geographically Weighted Regression (GWR) allows for the estimation of regression parameters specific to each location, enabling the parameters to vary spatially (Fotheringham and Oshan, 2016). GWR incorporates the geographic context and considers the local spatial variations, providing a more nuanced understanding of the relationships between variables compared to global regression models. The GWR is represented by the following equation:

$$y(TFR)_j = \beta_0(lon_j, lat_j) + \sum_{m=1}^{N} \beta_m\,(lon_j, lat_j)x_{mj} + \varepsilon_j$$

where j denotes district, j = 1, 2,…,981, and m denotes the number of selected explanatory variables. y_j denotes TFR at district j and (lon_j, lat_j) denotes the longitude and latitude coordinates of district j. β_m is a local estimated parameter of an explanatory variable, m. X_{mj} denotes the values of the m^{th} selected explanatory variable and ε_j refers a random error term. The location-specific parameter estimates allow the relationships between explanatory variables and TFRs to vary between districts. Parameter estimates for each explanatory variable at each district can be calculated in a matrix form:

$$\hat{\beta}(lon_j, lat_j) = [\mathbf{X}^T\mathbf{W}(lon_j, lat_j)\mathbf{X}]^{-1}[\mathbf{X}^T\mathbf{W}(lon_j, lat_j)\mathbf{Y}],$$

Thus, $\hat{\beta}(lon_j, lat_j)$varies with the values of $\mathbf{W}(lon_j, lat_j)$. For spatial weights, I adopted Gaussian weights and the bi-square weighting function. The $\mathbf{W}(lon_j, lat_j)$matrix contains the spatial weights in diagonal and 0 in its off-diagonal elements as:

$$\mathbf{W}(lon_j, lat_j) = \begin{bmatrix} w_1(lon_j, lat_j) & 0 & 0 \\ 0 & \ldots & 0 \\ 0 & 0 & w_n(lon_j, lat_j) \end{bmatrix}$$

By applying specific distance-based weighting functions, the estimation is achieved by using data from nearby locations. As a result, data from districts that are closer to the regression point are given bigger weights than data from districts that are farther away (Fotheringham et al., 2003). In this study I applied the commonly used adaptive bi-square kernel function as the distance-based weighting function (Brunsdon et al., 1998, Brunsdon et al., 1996):

$$W_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{b}\right)^2\right]^2 & , d_{ij} < d_{max} \\ 0 & , otherwise \end{cases}$$

where W_{ij} is the weight between district *i* and district *j*, d_{ij} is the distance between district *i* and district *ju*, and the bandwidth size, *b*, is maximum distance from regression location, *i*, which defines how many neighbouring observations should be included in the matrix. The optimal number of nearest neighbours was determined by minimising the corrected Akaike information criterion (Fotheringham et al., 2003). The advantage of using the GWR model is that it allows for the investigation of spatial variation in the correlation between the outcome variable and explanatory variables.

5.6. Results

5.6.1. Spatial Autocorrelation of total fertility rates at the DHS PSU and the district levels.

I found that there is evidence of spatial autocorrelation of TFR at the EDHS PSU level between 2000 and 2016 (Figure 5.2A), as it depicts that the empirical semi-variogram against the separation distance levels out as the separation distance increases (Figure 5.2B and Figure 5.2C). Our results further show that the spatial linear regression model outperforms the non-spatial linear regression model in modelling district-level TFRs in Ethiopia across four EDHS. The SLM reveals strong evidence of spatial correlations between district-level fertility and the fertility of neighbouring districts. This is evident from the positive spatial lag coefficients (ρ) in the SLM, ranging from 0.594 to 0.873, and their highly significant p-values ($p < 0.001$) across the four EDHS. Thus, high fertility in one district tends to be associated with high fertility in neighbouring districts (Table 5.2).

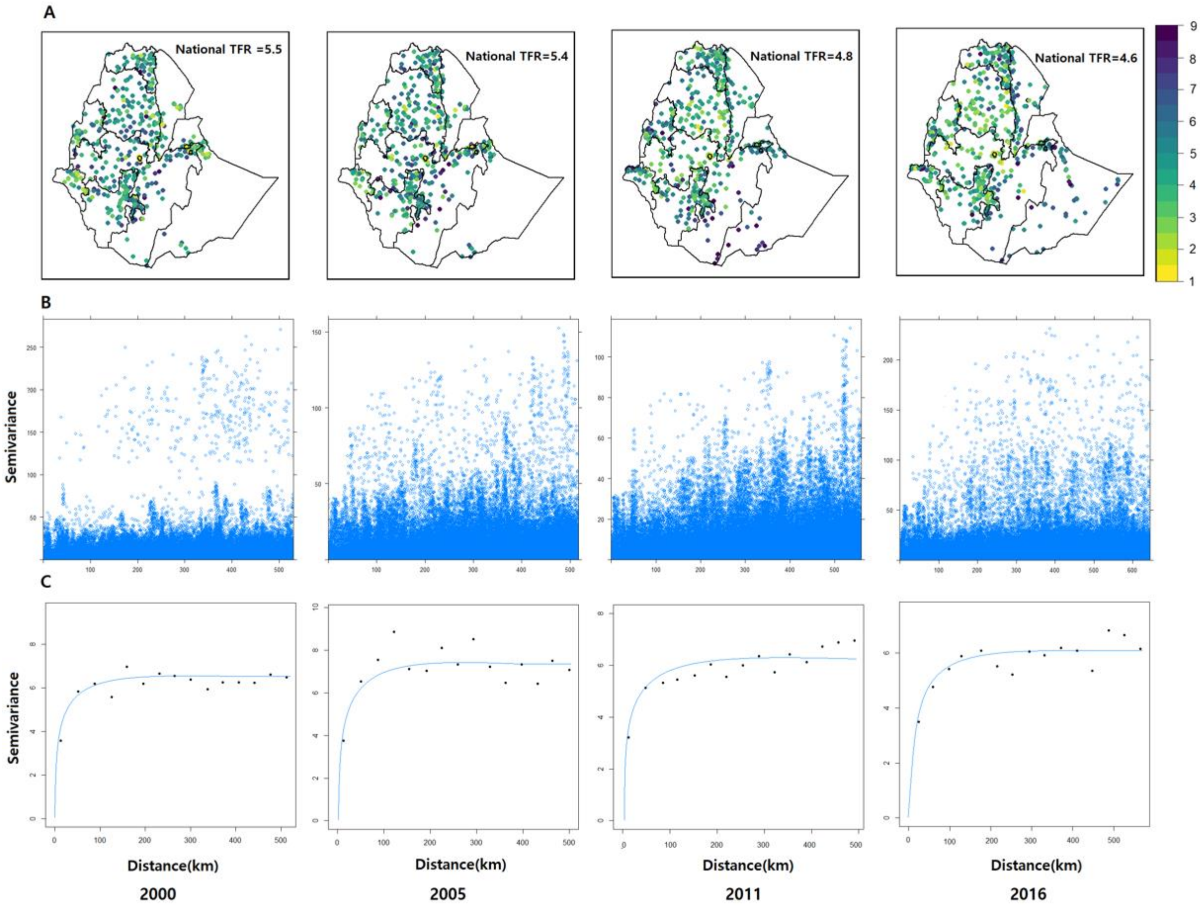


Figure 5.2. A. Estimated total fertility rate (TFR) at the primary sampling units (PSU) level and B. Variogram cloud and C. Empirical and exponential semi-variogram models of PSU-level TFRs, 2000-2016.

In all four years, the selected key proximate determinants, namely modern contraceptive prevalence (mCP) and median age at first marriage, as well as the selected socioeconomic factors, proportion of women living in urban areas and proportion of women with secondary and higher education, are consistently and negatively associated with district-level TFR in both the non-spatial and spatial linear regression models (p-value < 0.05). This implies that districts with higher mCP and median age at first marriage, as well as higher proportions of women living in urban areas and having secondary and higher education, tend to have lower district-level fertility.

Regarding the index of ethnolinguistic diversity, the results show a significant and positive association with district-level fertility in 2000, 2005, and 2011, in both the non-spatial and spatial linear regression models. This suggests that higher ethnolinguistic diversity is associated with higher district-level fertility in these three surveys. However, the results from the two models in 2016 are inconsistent. The SLM indicates a significant and positive association between the ethnolinguistic diversity index and district-level TFR, while the non-spatial regression model shows no significance. This suggests that the influence of ethnolinguistic diversity was underestimated by the non-spatial regression model in 2016. After controlling for the spatial lag of TFRs, the association between the ethnolinguistic diversity index and district-level TFR becomes significant at the 0.05 level.

Furthermore, a comparison of regression residuals between the non-spatial and spatial regression models for the four years (Figure 5.4) reveals that in 2000 and 2005, the residuals from both models exhibit similar spatially random patterns. However, in 2011 and 2016, spatial clustering of residuals is observed in the non-spatial model, while it is less pronounced in the SLM. This suggests that the spatial correlation effects of district-level fertility should be considered when investigating geographical variations in fertility in Ethiopia, particularly in recent years.

Although the SLM provides a better fit than the non-spatial linear regression model based on the AICc diagnostic criteria for the four EDHS, the global model fails to capture any spatial variations that may exist in the relationships between the outcome and explanatory variables across the study area. Therefore, a spatially non-stationary local modelling approach, namely the geographically weighted regression model, is adopted in the subsequent analysis.

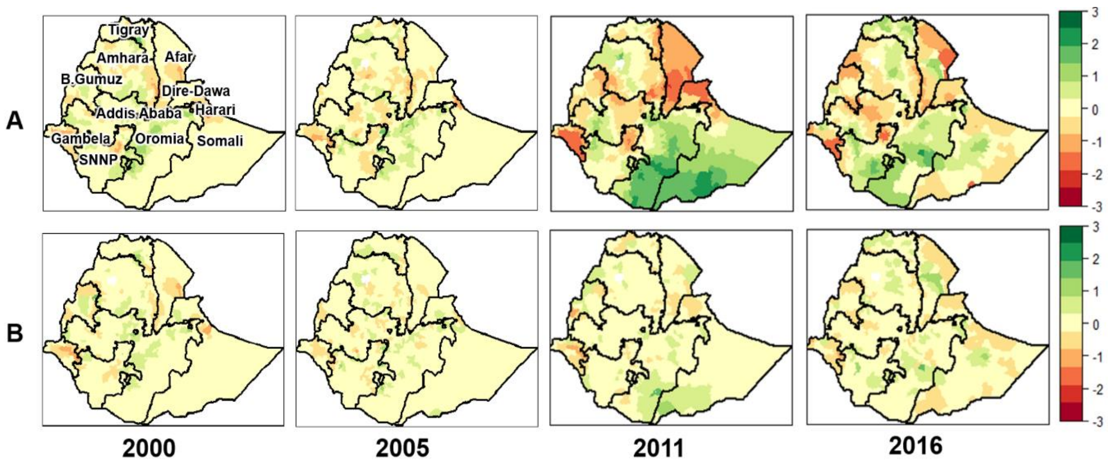


Figure 5.3. Regression residuals from A. non-spatial regression model and B. spatial regression model from 2000, 2005, 2011, and 2016

Table 5.2. Summary statistics of global and local regression models 2000-2016.

2000	Global model				Local model		
	Non-spatial linear model		Spatial lag model		GWR		
	Coefficient (95% CI)	P-value	Coefficient (95% CI)	P-value	Coefficient		
					Min	Median	Max
Explanatory variables							
mCP	-0.034 (-0.049 to -0.017)	< 0.001	-0.029 (-0.041to -0.018)	< 0.001	-0.579	-0.002	0.396
Age at first marriage	-0.051 (-0.072 to -0.029)	< 0.001	-0.009 (-0.024 to 0.001)	0.003	-1.529	-0.133	1.220
% female urban pop	-0.011 (-0.019 to -0.004)	0.001	-0.014 (-0.019 to -0.009)	< 0.001	-0.281	-0.018	0.207
% female education	-0.042 (-0.059 to -0.025)	< 0.001	-0.052 (-0.063 to -0.039)	< 0.001	-0.645	-0.039	0.537
Ethnic diversity	0.108 (0.011 to 0.207)	0.029	0.160 (0.093 to 0.229)	< 0.001	-0.895	-0.100	0.466
Spatial lag p			0.600 (0.566 to 0.634)	<0.001			
MODEL ASSESSMENT							
AIC	1029.264		394.415		-822.354		

2005	Global model				Local model		
	Non-spatial linear model		Spatial lag model		GWR		
	Coefficient (95% CI)	P-value	Coefficient (95% CI)	P-value	Coefficient		
					Min	Median	Max
Explanatory variables							
mCP	-0.013 (-0.019 to -0.006)	< 0.001	-0.015 (-0.020 to -0.010)	< 0.001	-0.259	-0.018	0.248
Age at first marriage	-0.076 (-0.084 to -0.044)	< 0.001	-0.049 (-0.066 to -0.032)	< 0.001	-1.425	-0.001	0.510
% female urban pop	-0.016 (-0.022 to -0.009)	< 0.001	-0.022 (-0.026 to -0.017)	< 0.001	-0.402	-0.018	0.301
% female education	-0.045 (-0.061 to -0.029)	< 0.001	-0.017 (-0.029 to -0.006)	0.003	-0.302	-0.034	0.424
Ethnic diversity	0.119 (0.040 to 0.198)	0.003	0.030 (-0.027 to 0.088)	0.303	-0.989	-0.046	0.857
Spatial lag p			0.595 (0.556 to 0.633)	<0.001			
MODEL ASSESSMENT							
AIC	881.321		354.486		-556.475		

2011	Global model				Local model		
	Non-spatial linear model		Spatial lag model		GWR		
	Coefficient (95% CI)	P-value	Coefficient (95% CI)	P-value	Coefficient		
					Min	Median	Max
Explanatory variables							
mCP	-0.089 (-0.096 to -0.083)	< 0.001	-0.013 (-0.016 to -0.010)	< 0.001	-0.524	-0.053	0.049
Age at first marriage	-0.099 (-0.159 to -0.018)	0.001	-0.013 (-0.034 to -0.008)	0.036	-1.140	-0.411	0.415
% female urban pop	-0.033 (-0.043 to -0.023)	< 0.001	-0.021 (-0.025 to -0.017)	< 0.001	-0.609	-0.017	0.054
% female education	-0.022 (-0.055 to -0.010)	0.018	-0.030 (-0.042 to -0.018)	< 0.001	-0.248	-0.045	0.094
Ethnic diversity	0.775 (0.567 to 0.982)	< 0.001	0.233 (0.157 to 0.310)	< 0.001	-0.821	-0.119	0.430
Spatial lag p			0.874 (0.849 to 0.898)	<0.001			
MODEL ASSESSMENT							
AIC	2568.963		782.787		-549.976		

2016	Global model				Local model		
	Non-spatial linear model		Spatial lag model		GWR		
	Coefficient (95% CI)	P-value	Coefficient (95% CI)	P-value	Coefficient		
					Min	Median	Max
Explanatory variables							
mCP	-0.060 (-0.063 to -0.057)	< 0.001	-0.016 (-0.018 to -0.013)	< 0.001	-0.617	-0.055	0.035
Age at first marriage	-0.156 (-0.205 to -0.107)	< 0.001	-0.039 (-0.064 to -0.013)	0.008	-1.158	-0.148	0.168
% female urban pop	-0.005 (-0.009 to -0.001)	0.028	-0.013 (-0.016 to -0.011)	< 0.001	-0.107	-0.002	0.095
% female education	-0.038 (-0.051 to -0.025)	< 0.001	-0.003 (-0.004 to -0.001)	<0.001	-0.436	-0.037	0.378
Ethnic diversity	0.065 (-0.202 to 0.048)	0.226	0.130 (0.064 to 0.195)	< 0.001	-0.625	-0.025	0.186
Spatial lag p			0.781 (0.747 to 0.816)	<0.001			
MODEL ASSESSMENT							
AIC	1794.018		677.950		-500.577		

5.6.3. Spatial heterogeneity of total fertility rates at the district level

To explore the local spatial variations in the relationships with the district-level TFR, the GWR was applied to the same set of explanatory variables used in the global models (Table 5.2). Then, we generated maps that represented the spatial distribution of their coefficients between 2000 and 2016 (Figure 5.5).

mCP had negative associations with district-level TFRs in most districts, and the negative coefficient values remained relatively consistent across the 981 districts in 2011 and 2016. This suggests that changes in mCP had similar effects on district-level TFRs in the majority of districts during these years. However, it should be noted that some districts in Tigray, Gambella, SNNP, and Somali regions showed positive coefficient estimates in the 2000 EDHS, as well as some districts in the Afar region in 2005 (Figure 5.5A). It is important to consider that mCP was generally very low in most districts in 2000 and 2005.

The negative coefficient estimates of median age at first marriage were found in most districts from each of the four years (Figure 5.5B). While the GWR coefficient values did not vary significantly across the 981 districts in 2000 and 2005, stronger coefficient estimates were observed in specific districts in Amhara, western Oromia, Harari, and northern Somali regions in the 2011 and 2016. This indicates that a delay of one year in the median age at first marriage was associated with a much lower district-level TFR in these districts in 2011 and 2016.

The coefficient estimates of the proportion of women living in urban areas showed negative associations with district-level TFRs, and these associations did not vary significantly across the 981 districts between 2000 and 2016 (Figure 5.5C). This suggests that changes in district-level TFRs in relation to changes in the proportion of women living in urban areas were spatially similar across the 981 districts between 2000 and 2016.

The negative coefficient estimates of the proportion of women having secondary and higher education were observed in most districts across all four EDHS, indicating that an increase in the proportion of women with secondary and higher education is associated with lower district-level TFRs in most districts (Figure 5.5D). However, positive coefficient values were observed in some districts in southern parts of Ethiopia in the 2000 and 2011.

The GWR coefficient map for the index of ethnolinguistic diversity reveals variations in coefficient estimates across the 981 districts in each of the four EDHS surveys (Figure 5.5E). Positive coefficients are frequently observed in some districts within the Amhara, Benishangul-Gumuz, and Gambela regions, as well as in the southern parts of Ethiopia across all four surveys. Conversely, negative estimates consistently emerge in urban regions, including Addis Ababa, Dire-Dawa, and

Harari regions. It is worth noting that the level of ethnolinguistic diversity at the zonal level is higher in the Benishangul-Gumuz, Gambela regions, and the three urban regions (Addis Ababa, Dire-Dawa, and Harari). This finding suggests that, on the one hand, higher ethnolinguistic diversity in the Benishangul-Gumuz and Gambela regions is associated with increased district-level Total Fertility Rates (TFRs). On the other hand, ethnolinguistic diversity in urban regions is linked to lower district-level TFRs. Additionally, within ethnolinguistically homogeneous areas at the zonal level, lower district-level TFRs are associated with ethnolinguistic homogeneity in the Amhara region and the southern part of Ethiopia. Conversely, higher district-level TFRs are associated with ethnolinguistic homogeneity in areas such as Afar, central Oromia, and northern Somali regions.

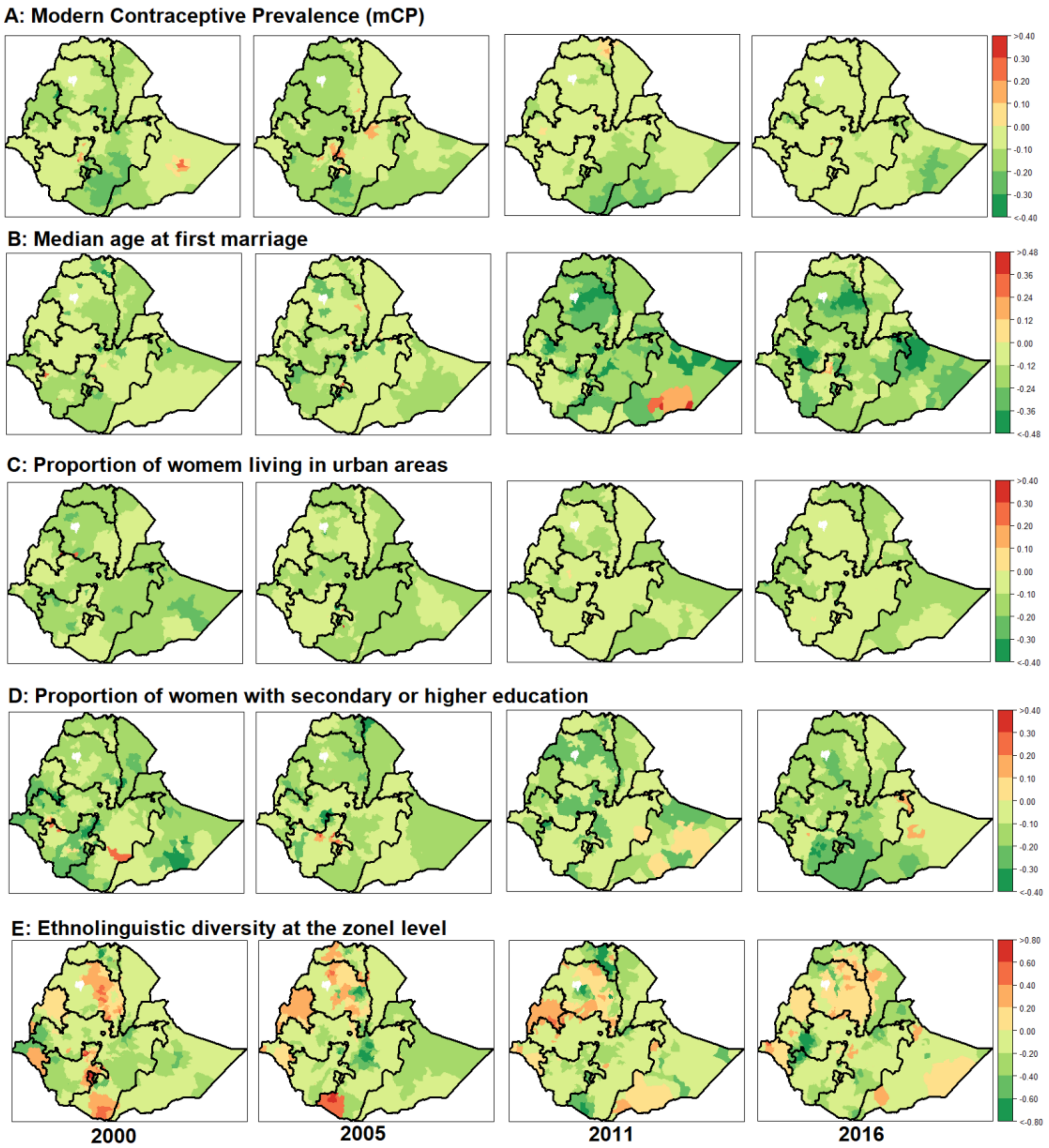


Figure 5.4. Local regression coefficients

A. modern contraceptive prevalence (mCP), **B**. median age at first marriage, **C**. proportion of women living in urban areas; **D**. Proportion of women with secondary or higher education; **E**. Index of ethnolinguistic diversity at the zonal level for 2000 – 2016 periods.

5.7. Discussion

Previous studies have highlighted the association between proximate and distal determinants and national or regional fertility levels in Ethiopia. However, these analyses conducted on large spatial scales have failed to capture the substantial geographical variations that exist at the district level. In Ethiopia, the district serves as a crucial administrative unit where health policies and programs are formulated and implemented through the Woreda Based Health Sector Planning (WBHSP). Despite the importance of districts, only a limited number of studies have examined how proximate and distal determinants are linked to district-level fertility. Furthermore, while there is growing evidence of spatial autocorrelation and spatial heterogeneity of fertility and its influencing factors in district-level analysis in HMICs, such spatial modelling of geographical variations in district-level fertility remains rare in sub-Saharan African countries. Hence, the overarching objective of this research was to investigate the spatial autocorrelation and heterogeneity of district-level fertility in relation to key selected proximate and distal determinants in Ethiopia.

The key findings of this study indicate that while different levels of proximate and distal determinants are crucial factors in explaining the geographical variations in district-level fertility, the fertility patterns at the district level, particularly in recent years, are significantly influenced by the spatial autocorrelation of district-level fertility. Moreover, spatial heterogeneity exists in the relationship between district-level fertility and both proximate and distal determinants in Ethiopia.

Firstly, this study shows that the spatial autocorrelation of total fertility rates (TFRs) persisted at both the DHS PSU and district levels between 2000 and 2016. The results obtained from semi-variogram analysis indicated a consistent presence of spatial autocorrelation of TFR at the DHS PSU level in all four years (2000, 2005, 2011, and 2016). Furthermore, the spatial autocorrelation coefficients derived from the SLM remained consistently positive and statistically significant throughout the four-year period, even after accounting for selected proximate and distal determinants of fertility. Moreover, the analysis of the spatial distribution of residuals obtained from the non-spatial linear model revealed clear evidence of spatial clustering in 2011 and 2016, while no such clustering was observed in the SLM. These findings suggest the presence of spatial autocorrelation in district-level fertility, particularly in recent years. It indicates that the fertility levels in a given district are increasingly influenced by the fertility levels of neighbouring districts, as well as the socioeconomic characteristics and reproductive behaviours within the district.

Secondly, geographical variations in district-level fertility across Ethiopia are not solely a result of geographical variations in the levels of key selected proximate and distal determinants of fertility. They also stem from spatially heterogeneous relationships between district-level fertility and both proximate and distal determinants across the 981 districts.

The GWR coefficient values of mCP and the two key socioeconomic factors, the proportions of women living in urban areas and having secondary or higher education, were consistently negative and relatively similar across the 981 districts in both 2011 and 2016. This implies that the geographical differences in mCP and the two key socioeconomic factors should be considered as significant indicators for explaining the geographical variations in district-level TFR in most districts in recent years. In other words, the decline in district-level TFR in most districts in recent years can be attributed to the increase in mCP, as well as the proportions of women living in urban areas and having secondary education. On one hand, the consistently lower fertility rates observed in urban districts in highly urbanized regions (Addis Ababa, Dire-dawa, and Harai regions) as discussed in Chapter 4, can be explained by the higher proportions of women living in urban areas and having secondary education. On the other hand, the recent fertility declines in rural districts of the Amhara region can be attributed particularly to the higher uptake of modern contraceptive use in the region. This suggests that the spread of lower fertility rates from Addis Ababa to the Amhara region, as observed in Chapter 4, is reasonable since Addis Ababa is the most urbanized area in Ethiopia and the Amhara region has reported a significant increase in the use of modern contraceptives due to the presence of family planning (FP) organizations and the government's focus on the region through the Health Extension Programme (HEP) since 2004 (Olson and Piller, 2013, Tegegne et al., 2020). These findings imply that rural and less educated women in the Amhara region are increasingly adopting modern methods of contraception, regardless of their socioeconomic conditions. This supports the view that well-designed FP programs can effectively reduce unintended births, as there is a significant level of unintended fertility in most societies in Sub-Saharan Africa (SSA) (Bongaarts, 1994). Advocates of FP programs further argue that intensive public FP campaigns can help reduce fears related to the side effects of modern family planning methods and potentially change the ideal number of children (Bongaarts and Bruce, 1995, Cleland et al., 2014). Ethiopia stands out as one of the few countries in SSA where the government shows a strong commitment to FP (Halperin, 2014, May and Rotenberg, 2020). While this study does not explore whether the national FP program in Ethiopia has reduced unintended births or desired family size, the results indicate that both the implementation of family planning programs and socioeconomic conditions are significant factors in explaining the spatial variations in district-level fertility in recent years in Ethiopia.

Furthermore, spatially heterogeneous relationships between district-level fertility and median age at first marriage were observed, particularly in 2011 and 2016. More specifically, stronger negative coefficient values were observed in certain districts within the Amhara region and the western parts of the Oromia region. This indicates that a one-year delay in first marriage is expected to have a greater impact on reducing TFR in these districts compared to other districts. It is worth noting that the Amhara region has the lowest median age at first marriage in the country, and there is a strong cultural link between marriage and childbearing, as evidenced by the low and nearly identical median ages at first marriage (15.7 years old) and first sexual intercourse (15.5 years old) in 2016 (Gage, 2013, Alem et al., 2020, Jones et al., 2018) (see appendix 4). To address early marriage in Ethiopia, the government revised the Criminal Code of 2005, imposing a maximum prison sentence of three years for marrying a girl aged between 13 and 17 years. The government also expressed its commitment to eradicating child marriage by 2025 in the National Strategy and Action Plan on Harmful Traditional Practices against Women in 2013 (MoWCYA, 2013, Abera et al., 2020). These national measures could have had a greater impact on districts within the Amhara regional state. It is known that delaying marriage acts as a strong facilitator of fertility decline in societies like the Amhara region, where childbearing outside of marriage is rare (Shapiro and Gebreselassie, 2014). Therefore, the stronger effects of delaying first marriage in the Amhara region may have partially contributed to the notable reduction in fertility observed in the region, as discussed in Chapter 4, in recent years. Moreover, the ethnolinguistic diversity had bidirectional effects on district-level fertility across different districts. Positive relationships were observed in two multi-ethnic regions (Benishagul-Gumuz and Gambela) and one ethnolinguistically homogeneous regional state (Amhara). On the other hand, negative relationships were particularly observed in multi-ethnic urban areas around Addis Ababa, Dire Dawa, and Harari regions.

The evidence of spatial autocorrelation of fertility and spatially heterogeneous relationships between fertility and key selected determinants in this study suggests that the spatial location and distance between districts increasingly play a role in explaining geographical variations in fertility in Ethiopia. Socioeconomic and cultural characteristics of districts in Ethiopia, even within the same region, significantly differ, and fertility in a district is influenced by its location, the characteristics of nearby districts, as well as its own characteristics. From a demographic theory perspective, both adaptationist and diffusionist approaches can jointly explain these spatial aspects of district-level fertility in Ethiopia. Firstly, the adaptationist approach argues that socioeconomic conditions are the best predictors of geographical variations in fertility. According to this approach, people tend to adapt to new socioeconomic conditions, resulting in fertility decline due to urbanization or increasing opportunity costs of childbearing. This study demonstrates that the coefficients for the proportions of women living in urban areas and having secondary education consistently had negative values in both non-spatial and spatial linear models. This evidence suggests that significant urban-rural differentiations in fertility in Ethiopia can be partially explained by varying levels of socioeconomic conditions between rural and urban districts. Secondly, the diffusionist approach argues that spatial diffusion or common cultural and linguistic interaction between neighboring geographic areas, regardless of socioeconomic conditions, plays a crucial role in geographical variations in fertility. Chapter 4 illustrated that the recent spatial spread of lower fertility from Addis Ababa to the Amhara region between 2011 and 2016 was similar to the spatial spread of higher mCP. Therefore, the spread of lower fertility in the Amhara region could be attributed to the more intensive family planning programs in the region, as previously explained. Additionally, it should be noted that the Amhara region is ethnolinguistically homogeneous, and Amharic, as the only working language of the federal government until 2020, is widely used in health policies (Ahmed and Seid, 2020). Consequently, within the diverse linguistic landscape of Ethiopia, this suggests that Amharic-speaking individuals in rural areas may have benefited from language-based reproductive health communication in the Amhara region, contributing to the adoption of modern contraception. Thus, ethnolinguistic diversity may be associated with higher fertility in multi-ethnic regional states like Gambela and Benishangul-Gumuz, supporting the idea that cultural or linguistic diversity can impede the diffusion of knowledge and attitudes favoring modern reproductive behaviors (De Broe and Hinde, 2006, Yüceşahin and Özgür, 2008, Bongaarts and Watkins, 1996). However, multi-ethnic urban areas around Addis Ababa, Dire Dawa, and the Harari region yielded contradictory results, as the index of ethnolinguistic diversity in these areas exhibited negative relationships with district-level fertility. From a diffusionist perspective, this is not a contradictory finding, as previous studies have shown that cultural diversification in urban areas often accelerates the diffusion of fertility changes, as communication networks in urban areas are more likely to transcend socioeconomic and cultural boundaries (Caldwell, 2006, Bongaarts and Watkins, 1996).

This study acknowledges several limitations that should be taken into account when interpreting the results. Firstly, the district-level estimates used in the analysis have associated uncertainty levels, which tend to be higher in certain regions due to smaller sample sizes and fewer number of Primary Sampling Units (PSUs). Therefore, caution should be exercised when interpreting the results, and the imprecise nature of the estimates should be considered when using them for practical purposes in health policy and planning at the district level. Although the approach used in this study aligns with DHS recommendations for geostatistical modelling and similar approaches have been employed by DHS themselves to produce health indicator estimates at smaller administrative boundaries (Mayala et al., 2019a, Janocha et al., 2021), the precision of the estimates for the 981 districts may still be limited. Secondly, the study employed cross-sectional analysis, which may not capture the temporal effects of proximate and distal determinants on fertility. Longitudinal data would provide a more comprehensive understanding of the dynamic relationships between these determinants and fertility over time.

Overall, it is important to recognise that geographical variations in district-level fertility in Ethiopia should be examined in the context of possible spatial autocorrelation of fertility and spatial heterogeneity in the relationships between fertility and both proximate and distal determinants. The district level is considered the most appropriate spatial scale in Ethiopia, given its significance in the country's national health policy planning and implementation. The findings highlight the spatial correlation of district-level TFRs and emphasise the importance of considering not only the local context but also the contexts of neighbouring districts when planning at the woreda (district) level. Furthermore, the strong association between mCP and geographical variations in district-level fertility, regardless of socioeconomic conditions, suggests the need for more district-specific FP programs tailored to the specific needs and characteristics of each district in Ethiopia. Such tailored programs can potentially improve the reproductive health status and outcomes in different districts.

5.8. Conclusion

This study reveals the presence of spatial autocorrelation and heterogeneity in fertility patterns concerning both proximate and distal determinants at the district level in Ethiopia in recent years. Consequently, any models that fail to account for spatial autocorrelations when explaining geographical variations in fertility in Ethiopia may yield biased results. Furthermore, the study highlights the increasing importance of space and place in understanding the recent geographical variations in fertility in Ethiopia. It is observed that socioeconomic and cultural characteristics vary among districts, even within the same region, and that fertility in a district is influenced not only by its own characteristics (adaptation effects) but also by the spatial location of the district and the characteristics of neighbouring districts (diffusion effects). Thus, this study offers additional insights into how the spatial location, distance between districts, socioeconomic and cultural characteristics, and reproductive behaviours within districts collectively contribute to shaping the geographical variations in district-level fertility in Ethiopia.

5.9. References

ABERA, M., NEGA, A., TEFERA, Y. & GELAGAY, A. A. 2020. Early marriage and women's empowerment: the case of child-brides in Amhara National Regional State, Ethiopia. *BMC international health and human rights,* 20**,** 1-16.

AHMED, M. & SEID, A. 2020. Association between exposure to mass media family planning messages and utilization of modern contraceptive among urban and rural youth women in Ethiopia. *International Journal of Women's Health,* 12**,** 719.

ALEM, A. Z., YESHAW, Y., KEBEDE, S. A., LIYEW, A. M., TESEMA, G. A., AGEGNEHU, C. D. & TESHALE, A. B. 2020. Spatial Distribution and Determinants of Early Marriage among Married Women in Ethiopia: A spatial and Multilevel Analysis. *BMC Women's Health,* 20**,** 1-13.

ANSELIN, L. & ARRIBAS-BEL, D. 2013. Spatial fixed effects and spatial dependence in a single cross-section. *Papers in Regional Science,* 92**,** 3-17.

BLANGIARDO, M. & CAMELETTI, M. 2015. *Spatial and spatio-temporal Bayesian models with R-INLA,* John Wiley & Sons.

BøCK, S., IMMITZER, M. & ATZBERGER, C. 2017. On the objectivity of the objective function—Problems with unsupervised segmentation evaluation based on global score and a possible remedy. *Remote Sensing,* 9**,** 769.

BONGAARTS, J. 1978. A framework for analyzing the proximate determinants of fertility.

BONGAARTS, J. 1994. The impact of population policies: Comment. *Population and Development Review,* 20**,** 616-620.

BONGAARTS, J. & BRUCE, J. 1995. The causes of unmet need for contraception and the social content of services. *Studies in family planning***,** 57-75.

BONGAARTS, J. & POTTER, R. E. 1983. *Fertility, biology, and behavior: An analysis of the proximate determinants,* Academic press.

BONGAARTS, J. & WATKINS, S. C. 1996. Social interactions and contemporary fertility transitions. *Population and development review***,** 639-682.

BRUNSDON, C., FOTHERINGHAM, A. S. & CHARLTON, M. E. 1996. Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis,* 28**,** 281-298.

BRUNSDON, C., FOTHERINGHAM, S. & CHARLTON, M. 1998. Geographically weighted regression. *Journal of the Royal Statistical Society: Series D (The Statistician),* 47**,** 431-443.

CALDWELL, J. C. 2006. The globalization of fertility behavior. *Demographic Transition Theory.* Springer.

CAMPISI, N., KULU, H., MIKOLAI, J., KLþSENER, S. & MYRSKYLæ, M. 2020. Spatial variation in fertility across Europe: Patterns and determinants. *Population, Space and Place***,** e2308.

CARLSSON, G. 1966. The decline of fertility: innovation or adjustment process. *Population studies,* 20**,** 149-174.

CLELAND, J., HARBISON, S. & SHAH, I. H. 2014. Unmet need for contraception: issues and challenges. *Studies in family planning,* 45**,** 105-122.

COSTA, R., BOCQUIER, P. & EGGERICKX, T. 2021. The Contribution of Diffusion to the Fertility Transition in Belgium (1887–1934). *Population and Development Review*.

DAVIS, K. & BLAKE, J. 1956. Social structure and fertility: An analytic framework. *Economic development and cultural change,* 4**,** 211-235.

DE BROE, S. & HINDE, A. 2006. Diversity in fertility patterns in Guatemala. *Population, Space and Place,* 12**,** 435-459.

ERULKAR, A. 2013. Early marriage, marital relations and intimate partner violence in Ethiopia. *International perspectives on sexual and reproductive health***,** 6-13.

EVANS, A. & GRAY, E. 2018. Modelling variation in fertility rates using geographically weighted regression. *Spatial Demography,* 6**,** 121-140.

FISH, T. D., JANOCHA, B., DONTAMSETTI, T. & MAYALA, B. K. 2020. Geospatial covariates: proxies for mapping urban-related indicators. *DHS Spatial Analysis Reports No. 19.* Rockville, Maryland, USA: ICF.
FOTHERINGHAM, A. S., BRUNSDON, C. & CHARLTON, M. 2003. *Geographically weighted regression: the analysis of spatially varying relationships*, John Wiley & Sons.
FOTHERINGHAM, A. S. & OSHAN, T. M. 2016. Geographically weighted regression and multicollinearity: dispelling the myth. *Journal of Geographical Systems,* 18**,** 303-329.
GAGE, A. J. 2013. Association of child marriage with suicidal thoughts and attempts among adolescent girls in Ethiopia. *Journal of Adolescent Health,* 52**,** 654-656.
GOLDSTEIN, J. R. & KLþSENER, S. 2014. Spatial analysis of the causes of fertility decline in Prussia. *Population and Development Review,* 40**,** 497-525.
GURMU, E. & ETANA, D. 2014. Age at first marriage and first birth interval in Ethiopia: analysis of the roles of social and demographic factors. *African Population Studies,* 28**,** 1332-1344.
HALPERIN, D. T. 2014. Scaling up of family planning in low-income countries: lessons from Ethiopia. *The Lancet,* 383**,** 1264-1267.
HAQUE, I., DAS, D. & PATEL, P. P. 2019. Reading the geography of India's district-level fertility differentials: a spatial econometric approach. *Journal of Biosocial Science,* 51**,** 745-774.
HOGAN, D. P. & BIRATU, B. 2004. Social identity and community effects on contraceptive use and intentions in southern Ethiopia. *Studies in Family Planning,* 35**,** 79-90.
ICF International and Central Statistical Agency. Ethiopia DHS 2000. Addis Ababa, Ethiopia and Calverton, MD: Central Statistical Agency, ICF International, 2000
ICF International and Central Statistical Agency. Ethiopia DHS 2005. Addis Ababa, Ethiopia and Calverton, MD: Central Statistical Agency, ICF International, 2005.
ICF International and Kenya National Bureau of Statistics. Kenya DHS 2014. Nairobi, Kenya and Calverton, MD: Kenya National Bureau of Statistics, ICF International, 2014.
ICF International and Central Statistical Agency. Ethiopia DHS 2016. Addis Ababa, Ethiopia and Calverton, MD: Central Statistical Agency, ICF International, 2016.
ICF International and Central Statistical Agency. 2015. *The DHS Program STATcompiler* [Online]. USAID. Available: http://www.statcompiler.com [Accessed December 19 2021].
JANOCHA, B., DONOHUE, R. E., FISH, T. D., MAYALA, B. K. & CROFT, T. N. 2021. Guidance and recommendations for the use of indicator estimates at subnational administrative level 2. *DHS Spatial Analysis Report 20.* Rockville, Maryland, USA: ICF.
JONES, N., TEFERA, B., EMIRIE, G. & PRESLER-MARSHALL, E. 2018. 'Sticky'gendered norms: Change and stasis in the patterning of child marriage in Amhara, Ethiopia. *Empowering Adolescent Girls in Developing Countries.* Routledge.
KOSTOV, P. 2010. Model boosting for spatial weighting matrix selection in spatial lag models. *Environment and Planning B: Planning and Design,* 37**,** 533-549.
LAELAGO, T., HABTU, Y. & YOHANNES, S. 2019. Proximate determinants of fertility in Ethiopia; an application of revised Bongaarts model. *Reproductive health,* 16**,** 13.
MAY, J. F. & ROTENBERG, S. 2020. A Call for Better Integrated Policies to Accelerate the Fertility Decline in Sub-Saharan Africa. *Studies in Family Planning,* 51**,** 193-204.
MAYALA, B. K., DONTAMSETTI, T., FISH, T. D. & CROFT, T. 2019a. *Interpolation of DHS survey data at subnational administrative Level 2*, DHS Program, ICF.
MAYALA, B. K., DONTAMSETTI, T., FISH, T. D. & CROFT, T. N. 2019b. Interpolation of DHS survey data at subnational administrative Level 2. *DHS Spatial Analysis Reports No. 17.* Rockville, Maryland, USA: ICF.
MEKONNEN, Y., TELAKE, D. S. & WOLDE, E. 2018. Adolescent childbearing trends and sub-national variations in Ethiopia: a pooled analysis of data from six surveys. *BMC pregnancy and childbirth,* 18**,** 1-13.
MOH 2021. Health Sector Transformation Plan (2020/21-2024/25). Federal Ministry of Health Addis Ababa, Ethiopia.
MOWCYA 2013. *National strategy and action plan on harmful traditional practices (HTPs) against women and children in ethiopia federal democratic republic of ethiopia ministry of women (MoWCYA),* Addis Ababa, Ministry of Women, Children and Youth Affairs.

OLSON, D. J. & PILLER, A. 2013. Ethiopia: an emerging family planning success story. *Studies in Family Planning,* 44**,** 445-459.
REDA, A. A. & LINDSTROM, D. 2014. Recent trends in the timing of first sex and marriage among young women in Ethiopia. *Etude de la population africaine= African population studies,* 28**,** 1157.
ROGERS, E. & STEPHENSON, R. 2018. Examining temporal shifts in the proximate determinants of fertility in low-and middle-income countries. *Journal of biosocial science,* 50**,** 551-568.
RUE, H., MARTINO, S. & CHOPIN, N. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology),* 71**,** 319-392.
SABATER, A. & GRAHAM, E. 2019. International migration and fertility variation in Spain during the economic recession: A spatial Durbin approach. *Applied Spatial Analysis and Policy,* 12**,** 515-546.
SALVATI, L., BENASSI, F., MICCOLI, S., RABIEI-DASTJERDI, H. & MATTHEWS, S. A. 2020. Spatial variability of total fertility rate and crude birth rate in a low-fertility country: patterns and trends in regional and local scale heterogeneity across Italy, 2002–2018. *Applied Geography,* 124**,** 102321.
SHAPIRO, D. & GEBRESELASSIE, T. 2014. Marriage in sub-Saharan Africa: Trends, determinants, and consequences. *Population Research and Policy Review,* 33**,** 229-255.
SHIFERAW, S., ABDULLAH, M., MEKONNEN, Y., MAñGA, A., AKINYEMI, A., AMOUZOU, A., FRIEDMAN, H., BARROS, A. J. & HOUNTON, S. 2015. Trends in contraceptive use and distribution of births with demographic risk factors in Ethiopia: a sub-national analysis. *Global health action,* 8**,** 29720.
SINGH, A., KUMAR, K., PATHAK, P. K., CHAUHAN, R. K. & BANERJEE, A. 2017. Spatial patterns and determinants of fertility in India. *Population,* 72**,** 505-526.
STAVETEIG, S., GEBRESELASSIE, T. & KAMPA, K. T. 2018. Absolute poverty, fertility preferences, and family planning use in FP2020 focus countries. *DHS Comparative Reports No. 48.* Rockville, Maryland, USA: ICF.
TEGEGNE, T. K., CHOJENTA, C., FORDER, P. M., GETACHEW, T., SMITH, R. & LOXTON, D. 2020. Spatial variations and associated factors of modern contraceptive use in Ethiopia: a spatial and multilevel analysis. *BMJ open,* 10**,** e037532.
TESHOME, S. B. & HOEBINK, P. 2018. Aid, ownership, and coordination in the health sector in Ethiopia. *Development Studies Research,* 5**,** S40-S55.
TESSEMA, Z. T. & TAMIRAT, K. S. 2020. Determinants of high-risk fertility behavior among reproductive-age women in Ethiopia using the recent Ethiopian Demographic Health Survey: a multilevel analysis. *Tropical Medicine and Health,* 48**,** 1-9.
UNICEF 2018. Ending child marriage: a profile of progress in Ethiopia. *New York: UNICEF*.
VITALI, A. & BILLARI, F. C. 2017. Changing determinants of low fertility and diffusion: A spatial analysis for Italy. *Population, Space and Place,* 23**,** e1998.
WANG, D. & CHI, G. 2017. Different Places, Different Stories: A Study of Spatial Heterogeneity of County-Level Fertility in China. *Demographic research,* 37**,** 493.
WARD, M. D. & GLEDITSCH, K. S. 2018. *Spatial regression models*, Sage Publications.
YþCEŞAHIN, M. M. & ÖZGþR, E. M. 2008. Regional fertility differences in Turkey: persistent high fertility in the southeast. *Population, Space and Place,* 14**,** 135-158.

Chapter 6

Discussion and Conclusions

6. Chapter 6: Discussion and conclusions

6.1. Overview

In the previous chapters, I have described and explored spatial aspects of geographical variations in district-level fertility associated with key selected proximate and distal determinants in Ethiopia. In chapter 6, I review and synthesise the key findings of the DrPH research and discuss the strengths and limitations of this study. Then, I will also discuss implications in Ethiopia, and possible directions for future work beyond the DrPH.

6.2. Summary of findings

The main aim of this thesis was to describe and explore spatial aspects of geographical variations in district-level fertility in association with key selected proximate and distal determinants in Ethiopia between 2000 and 2016. This aim was addressed by the following objectives:

1. Objective 1: To estimate TFRs and key selected proximate and distal determinants for 981 districts in 2000, 2005, 2011 and 2016 by using a geostatistical modelling approach (Chapter 4).

2. Objective 2: To describe and explore spatial and temporal patterns of TFR and key selected proximate and distal determinants at the district level in 2000, 2005, 2011 and 2016 (Chapter 4).

3. Objective 3: To assess effects of key selected proximate and distal determinants on geographical variations in fertility at the district level between 2000-2016 with a non-spatial model (Chapters 5).

4. Objective 4: To assess spatial dependency of district-level fertility by using a spatial method (Chapters 5).

5. Objective 5: To explore spatial heterogeneity in relationships between TFRs and both proximate and distal determinants in Ethiopia by using geographically weighted regression between 2000 and 2016 (Chapters 5).

In this section, I will summarise the main findings of the DrPH in relation to these five study objectives, and discuss how the findings help explain fertility changes at the district-level in Ethiopia.

6.2.1. Objective 1: To estimate total fertility rates and key selected proximate and distal determinants for 981 districts in 2000, 2005, 2011 and 2016 by using a geostatistical modelling approach.

The key findings from Objective 1 of the study indicate that there have been increasing differences in Total Fertility Rates (TFRs) between regions in recent years. Additionally, district-level TFRs have become more distinct from their respective median regional TFRs. Similar trends were observed in modern contraceptive prevalence (mCP) (Figure 4.2. & Figure 4.3).

The three urban regions (Addis Ababa, Dire Dawa, and Harari) consistently exhibited lower regional-level fertility compared to the other eight regions. The district-level TFRs in these urban regions remained similar to their respective median regional-level TFRs between 2000 and 2016. In contrast, in 2000 and 2005, the district-level TFRs within other eight regions aligned with their respective median regional-level TFRs, except for the Oromia region. However, in 2011 and 2016, the median regional-level TFRs among the eight regions began to diverge, and the district-level TFRs within these regions became increasingly different from their respective median regional-level TFRs.

Specifically, Oromia, Amhara, Tigray, and SNNP regions experienced declines in their regional-level TFRs between 2000 and 2016, and the geographical variations in district-level fertility within these regions widened in recent years. The four Developing Regional States (DRS) exhibited different trends in regional-level fertility but displayed similar patterns of variation in district-level fertility. The two eastern and ethnically homogeneous regions (Afar and Somali) had higher or increasing regional-level fertility between 2011 and 2016, and district-level TFRs within these regions started to vary. The two western and multi-ethnic regions (Gambela and Benishangul-Gumuz) experienced gradual declines in regional-level fertility, and the district-level TFRs within these regions showed variations, particularly in 2011 and 2016.

Regarding the selected proximate determinants, similar changes in mCP were observed between 2000 and 2016. In 2000 and 2005, the eight regional states (excluding the three urban regions) had relatively similar levels of mCP with minimal variations within regional states. However, in 2011 and 2016, mCP levels diverged both between and within regional states. In contrast, age at first marriage varied between regional states in 2000 and 2005 but gradually increased, leading to smaller variations between districts and regional states by 2016.

In terms of the selected distal determinants, regional and district-level variations remained relatively stable between 2000 and 2016 compared to the proximate determinants. The two socioeconomic determinants, the proportion of women living in urban areas and those with secondary or higher education, remained low (less than 30% and 20% for urban residency and secondary education, respectively, in 2016) and exhibited similar patterns among the eight regions (excluding the three urban regions) between 2000 and 2016. The ethnolinguistic index at the zone level was higher in the three urban regions (Addis Ababa, Dire Dawa, Harari) and the two western DRS (Gambela and Benishangul-Gumuz), and the district-level variations in the ethnolinguistic index remained relatively constant between 2000 and 2016.

While previous studies in Ethiopia have focused on exploring geographical variations in fertility and proximate and distal determinants between the country's eleven regions (Laelago et al., 2019, Tessema and Tamirat, 2020), the findings presented in Chapter 4 of this study go beyond these previous research efforts. This study examines geographical variations in district-level fertility and proximate and distal determinants within the eleven regions of Ethiopia between 2000 and 2016. The decision to estimate study variables at the third administrative unit, the districts (Admin 3), was driven by their significance in health planning and service delivery within Ethiopia's woreda-based health sector plan. To analyse the geographical distribution of TFR and selected proximate and distal determinants, this study proposed the application of model-based geostatistics using the INLA-SPDE model approach. This approach was utilized to assess the TFR and determinants at 981 districts in Ethiopia between 2000 and 2016.

The findings of this study are particularly important as they demonstrate that relying solely on national or regional data provides an inadequate description of recent geographical variations in fertility in Ethiopia. This inadequacy can pose challenges in planning for and monitoring decentralized national health programs at the district level in Ethiopia. Therefore, by focusing on district-level analysis, this study sheds light on the need to consider the specific characteristics and variations within districts, emphasizing the importance of district-level data in informing effective health planning and implementation.

6.2.2. Objective 2: To describe and explore spatial and temporal patterns of study variables in 2000, 2005, 2011 and 2016.

Two key findings from the study objective 2 are as follows: Firstly, there is a clear spatial spread of lower fertility from Addis Ababa to the northern and western parts of Ethiopia between 2000 and 2016. Secondly, the changes in the spatial distribution of fertility were more consistent with changes in the spatial distribution of proximate determinants, rather than distal determinants (Figure 4.4).

Firstly, in Chapter 4, it was evident that the changes in the spatial patterns of district-level fertility were not randomly distributed; they followed certain spatial patterns. While urban and multi-ethnic regions around Addis Ababa, Dire Dawa, and Harari consistently exhibited lower fertility rates compared to the other eight regions between 2000 and 2016, lower district-level fertility was predominantly restricted to these urban regions between 2000 and 2005. However, a notable observation was that lower district-level fertility appeared to extend from Addis Ababa to the northern and western parts of the country in 2011 and 2016.

Secondly, the changes in the spatial pattern of district-level fertility resembled the changes in the spatial patterns of the two selected proximate determinants, rather than the three selected distal determinants. Higher modern contraceptive prevalence (mCP) was particularly observed around the three urban and multi-ethnic regions between 2000 and 2005. However, in 2011, a spatial spread of higher mCP was observed from Addis Ababa to the western and northern parts of Ethiopia, including several rural districts in the Amhara regional state. Furthermore, the median age at first marriage was notably lower in the Amhara region compared to other parts of Ethiopia between 2000 and 2005. However, an increase in the median age at first marriage was observed, particularly in the Amhara region, between 2011 and 2016. In contrast, there were no apparent changes in the spatial distributions of the three selected distal determinants between 2000 and 2016.

As mentioned earlier, districts play a crucial role in health planning and service delivery in Ethiopia, as decentralized health services are provided at the district level through the "woreda-based annual health sector planning," which is a key strategy in the current Ethiopia Health Sector Transformation plan (2019/20-2024/25). However, due to the lack of district-level data, we have limited information about the geographical distribution of fertility and the key factors influencing fertility at the district level. This thesis clearly demonstrates that a spatial spread of fertility decline has occurred from Addis Ababa to the northern and western parts of Ethiopia in recent years, rather than fertility declines being evenly observed across the country. This finding is significant because it suggests that these observed spatial patterns may influence how the government, international organizations, and NGOs implement public health measures at the district level.

6.2.3. Objective 3: To assess effects of key selected distal and proximate determinants on geographical variations in fertility at the district level between 2000-2016 with a non-spatial model.

A key finding from the study objective 3 was that modern contraceptive prevalence (mCP), median age at first marriage, proportions of women living in urban areas, and those with secondary education were negatively associated with district-level fertility, whereas ethnolinguistic diversity had a positive influence on district-level fertility in the non-spatial model between 2000 and 2016 (Table 5.2).

I found that that mCP and median age at first marriage had generally negative relationships with district-level total fertility rates (TFRs) between 2000 and 2016. Previous studies in Ethiopia have shown significant spatial variations in mCP and child marriage (Ebrahim et al., 2021, Tegegne et al., 2020, Alem et al., 2020), and these two factors are important in explaining regional variations in fertility. This study further demonstrates that mCP and median age at first marriage are crucial indicators in explaining geographical variations in district-level fertility.

Additionally, the two socioeconomic determinants, the proportions of women living in urban areas and having more than secondary education, were also negatively related to district-level fertility. This suggests that increases in the proportions of these two factors are associated with lower district-level TFRs. According to the DHS STATcompiler (ICF, 2015), this finding aligns with the results from the DHS STATcompiler, which showed that Ethiopia had the largest fertility difference by residence and female educational level among sub-Saharan African countries surveyed by the DHS program (appendix 3). This result implies that geographical variations in socioeconomic conditions are reflected in geographical variations in district-level fertility, supporting the adaptationist view that fertility variation can be primarily seen as a reaction to socioeconomic conditions (Becker, 1960a, Easterlin, 1975).

Furthermore, the non-spatial linear regression showed that ethnolinguistic diversity is positively associated with district-level fertility between 2000 and 2016. This suggests that multi-ethnic areas are typically associated with higher fertility, while ethnically homogeneous areas are associated with lower fertility. This result aligns with the diffusionist approach, which suggests that fertility decline spreads across geographically neighboring areas, especially when they share a common language Ethiopia has over 90 ethnolinguistic groups, and ethnolinguistic identities are the most important criteria for shaping regional boundaries. Chapter 4 of the study presented the geographical distribution of ethnolinguistic diversity at the zonal level, which can indicate the extent to which districts are surrounded by similar or different ethnolinguistic districts. Apart from the three urban regions, Benishangul-Gumuz and Gambela regional states consistently had higher levels of ethnolinguistic diversity at the zone level, indicating that districts in these two regions are surrounded by different ethnolinguistic districts. This finding suggests that, after controlling for other relevant factors, ethnolinguistic diversity in Benishangul-Gumuz and Gambela may hinder the diffusion of new knowledge and attitudes favouring reproductive behaviour. On the other hand, common languages in ethnically homogeneous regions, such as the Amhara region, can facilitate the diffusion of the idea of smaller families.

Previous studies have examined the associations between proximate and distal determinants and fertility levels in Ethiopia (Laelago et al., 2019, Eyasu, 2015, Mengesha et al., 2018), but these studies often focused on national or regional levels. However, it is important to note that, apart from the three urban regions, the land areas of the eight regions are very large, with the land area of the

Oromia region (353,690 km²) being larger than that of the United Kingdom (243,610 km²). Therefore, district-level analysis is important since national- or regional-level analyses likely mask changes in district-level TFRs in association with key proximate and distal determinants of fertility.

6.2.4. Objective 4: To assess spatial autocorrelation of district-level fertility by using a spatial model

A key finding from study objective 4 is that spatial autocorrelation of district-level fertility has been an important aspect of geographic variations in district-level fertility in recent years.

I found that spatial autocorrelation of TFRs at the DHS PSU level were consistently observed from semi-variogram analysis between 2000 and 2016 (Figure 5.3). Additionally, the direction of correlation between district-level TFRs and key selected proximate and distal determinants in the Spatial Lag Model (SLM) was similar to the result from the non-spatial model. However, the spatial lag term (p) in the SLM was consistently positive and significant between 2000 and 2016, suggesting that district-TFRs are spatially correlated even after controlling for key selected proximate and distal determinants (Table 5.2). This thesis further showed that regression residuals from non-spatial and spatial linear regression models were relatively similar and not spatially clustered in 2000 and 2005. However, the residuals were much smaller, and the spatial cluster of residuals was less obvious in the SLM than that of the non-spatial linear regression model in 2011 and 2016 (Figure 5.4).

A number of recent studies in high and middle-income countries (HMICs) have called for attention to the existence of spatial autocorrelation of fertility at the local or district level (Salvati et al., 2020, Vitali and Billari, 2017, Campisi et al., 2020, Burillo et al., 2020). Although geographically referenced data have become increasingly available in sub-Saharan Africa (Pezzulo et al., 2021, Tessema et al., 2020, Benza et al., 2017), studies that model spatial autocorrelation in district-level fertility are rare in SSA.

This thesis clearly shows that spatial location and distance between districts are increasingly important aspects of fertility changes at the district level in recent years in Ethiopia. This finding suggests that TFR in a district can be affected by TFRs of nearby districts. This implies that any models that do not consider distance-based spatial autocorrelations in explaining geographical variations in fertility in Ethiopia may over- or underestimate the effects of key determinants of fertility on district-level fertility, especially in recent years.

6.2.5. Objective 5: To explore spatial heterogeneity in relationships between TFRs and both proximate and distal determinants in Ethiopia by using geographically weighted regression between 2000 and 2016.

A key finding from study objective 5 is that spatially heterogeneous relationships between ethno-linguistic diversity and district-level fertility were observed in terms of direction, whereas there was a relatively small degree of spatial heterogeneity in the relationships between district-level fertility and mCP, age at first marriage, the proportion of women living in urban areas, and having secondary education (Figure 5.5).

Although non-spatial and spatial regression models generally showed that ethnolinguistic diversity is positively related to district-level fertility in general, the Geographically Weighted Regression (GWR) showed that the relationships were spatially heterogeneous in terms of direction. Specifically, while multi-ethnic urban regions (Addis Ababa, Dire Dawa, and Harari) had a negative relationship between ethnolinguistic diversity and district-level fertility, other multi-ethnic regions (Benishangul-Gumuz and Gambela) had positive relationships.

Additionally, mCP, age at first marriage, the proportion of women living in urban areas, and having secondary education generally showed negative and spatially similar relationships with district-level fertility across 981 districts in recent years. This means that the influences of changes in these four variables on district-level fertility are relatively similar across 981 districts. In other words, geographical variations in the levels of mCP, age at first marriage, and the proportion of women in urban areas and having secondary education matter in geographical variations in district-level fertility, which supports the view that both socioeconomic conditions and family planning programs are strong predictors of geographical variations in fertility (Bongaarts and Bruce, 1995, Cleland et al., 2014). For age at first marriage, in particular, the GWR showed that the regression coefficients were spatially heterogeneous in terms of magnitude in 2011 and 2016. It shows that districts in the Amhara region had stronger negative effects, which means that a delay in first marriage by one year may decrease district-level fertility more in the Amhara region, which has the lowest median age at first marriage in the country.

Previous studies in high- and middle-income countries (HMICs) have demonstrated that spatially varying relationships between district-level fertility and socioeconomic variables do exist, and therefore a single, one-size-fits-all model may not easily summarize the spatial dimensions of fertility differentials and their underlying correlates (Wang and Chi, 2017, Haque et al., 2019). Such district-level fertility studies are rare in sub-Saharan Africa (SSA), but they are important in SSA because socioeconomic, linguistic, and cultural contexts substantially vary between districts in many SSA countries. This approach is particularly essential in Ethiopia, where geographical boundaries largely reflect ethnolinguistic boundaries. Therefore, the geographically weighted regression model was used to investigate the spatially varying relationships between district-level fertility and key selected proximate and distal determinants in Ethiopia. The spatially varying relationship between ethnolinguistic diversity and district-level fertility, found in this thesis, implies that ethnolinguistic diversity facilitates fertility decline in some districts but prevents it in other districts.

6.2.6. Discussion of findings

Overall, this DrPH thesis presents three main findings. Firstly, geographical variations in district-level fertility have significantly emerged in recent years, and the emerging geographical variations in district-level fertility follows a certain spatial pattern where lower fertility has spread from Addis Ababa to the northern and western parts of the country. Secondly, recent spatial changes in district-level fertility were more consistent with changes in the two selected proximate determinants rather than changes in the levels of the three selected distal determinants. Thirdly, the geographical variations in district-level fertility in Ethiopia were associated with the spatial location of and distance between districts, as well as the socioeconomic and cultural characteristics, and reproductive behaviours within the districts.

1) Emergence of geographical variations in district-level fertility

First, this thesis showed that geographical variations in district-level fertility in Ethiopia have changed and widened in recent years. In 2000 and 2005, lower fertility levels were particularly observed around the three small and urban regions (Addis Ababa, Dire-Dawa and Harari). However, lower fertility levels were increasingly observed in non-urban districts as well as urban regions in 2011 and 2016, where the spatial spread of district-level fertility decline from the capital to northern and western parts of Ethiopia were clearly detected. Exploring the spatial spread of fertility decline beyond the simple urban and rural division at the district level is not a new concept. Previous studies in HMICs have also revealed the existence of spatial variations in fertility across district-level areas within a country, which are associated with subnational conditions such as cultural and spatial contexts, as well as urban-rural contexts (Wang and Chi, 2017, Vitali and Billari, 2017, Campisi et al., 2020, Haque et al., 2019, Costa et al., 2021). However, district-level studies in HMICs have relied on existing data, such as national census and vital statistics. One of the key challenges for achieving equity in health service provision in many SSA countries is the lack of district-level health and population data, which hampers the monitoring and evaluation of health outcomes across districts within a country. Moreover, official fertility and health statistics in SSA usually rely on national estimates, potentially masking underlying heterogeneity in fertility and health outcomes within countries. Therefore, the Sustainable Development Goals (SDGs) call for disaggregated data by geographic location and other relevant characteristics in national contexts to uncover subnational disparities and ensure that no one is left behind in the implementation of the SDGs. In Ethiopia, districts (Admin 3) are essential administrative units for health planning and service provision, explicitly recognised by Ethiopia's Woreda-Based Annual Health Sector Planning (WBHSP). Although previous studies have described the geographical distribution of selected health outcomes at the regional (Admin 1) or zone level (Admin 2) (Mayala et al., 2019a, Janocha et al., 2021), health policy and services are decentralized and implemented based on district boundaries in accordance with the WBHSP. Therefore, a description of geographical variations at the Admin 1 or 2 level provides inadequate information about variations in population and health outcomes at the district level to monitor national health programs such as the WBHSP. To address this gap, I utilised Ethiopia DHS data from 2000 to 2016 to provide geographical descriptions of district-level TFRs in Ethiopia using the Bayesian model-based framework, which is the main novelty of this study. This approach provides valuable information for other Sub-Saharan African (SSA) countries where significant district-level variations in population and health outcomes exist. The lack of district-level data hampers the monitoring of geographical differences in health status at the district level in these countries. This approach can contribute not only to reducing sub-national health disparities within a country but also to meeting the SDGs' call for the development of better disaggregated data by geographical areas.

2) Role of proximate determinants on geographical variations in district-level fertility

This thesis provides clear evidence that the recent changes in spatial variations in district-level fertility in Ethiopia were more closely aligned with the spatial changes in the levels of the two selected proximate determinants, rather than the spatial changes in the levels of the three selected distal determinants. This finding was determined by conducting an investigation of both proximate and distal determinants of fertility at the district level, aiming to describe the association between changes in geographical variations in district-level fertility and the key determinants of fertility.

Previous studies on district-level fertility in HMICs often focused primarily on distal determinants, such as socioeconomic and cultural variables, while neglecting the role of proximate determinants in understanding geographical variations in fertility. Regarding the three distal determinants, I observed overall increases in the proportions of women living in urban areas and with secondary or higher education across districts between 2000 and 2016. However, these increases were relatively small, except for consistently higher proportions in urban regions. Similarly, although there were variations in ethnolinguistic diversity at the zone level across eleven regions, I found minimal changes in the geography of ethnolinguistic diversity between 2000 and 2016. Nevertheless, the small changes in the spatial distributions of the two socioeconomic variables do not imply that they are no longer important factors in explaining the geographic variations in district-level fertility. Districts with lower fertility levels in urban areas consistently exhibited higher proportions of women living in urban areas and with secondary or higher education. Hence, the levels of these two selected socioeconomic variables strongly correlate with urban-rural differences in district-level fertility. This association aligns with the adaptationist view, which suggests that socioeconomic changes play a significant role in driving geographical variations in fertility (Bryant, 2007, Shapiro and Tenikue, 2017, Behrman, 2015).

Notably, differences in fertility levels among urban-rural residents and educational levels were largest compared to other countries surveyed by the DHS programmes (appendix 3). Furthermore, a recent study demonstrated that Addis Ababa had the lowest fertility rates among 932 first administrative units (Admin 1) in SSA (Pezzulo et al., 2021). This implies that residents in Addis Ababa are already accustomed to smaller families, and rural-to-urban migrants also tend to adapt their fertility behaviour quickly to the new socioeconomic conditions in their urban destinations. Alternatively, these migrants may selectively move into urban areas to capitalise on the greater income and job opportunities offered by cities (Kulu, 2005). In Ethiopia, approximately 80% of the population resides in rural areas, and over 70% of the population is estimated to be employed in the agricultural sector (OECD, 2020). Hence, rural-to-urban migration decisions in Ethiopia are likely driven by voluntary migration for better job or income opportunities, suggesting that rural-urban migrants are more inclined to adapt to new urban lifestyles rather than maintaining the higher fertility levels prevalent in their original districts (Gibson and Gurmu, 2012). However, despite the significance of adaptation effects, the relatively gradual or small changes in the levels of the two socioeconomic factors between 2000 and 2016 cannot adequately account for the disproportionate changes in fertility across districts, particularly in 2011 and 2016.

In contrast to the three distal determinants discussed earlier, recent changes in the geography of the two selected proximate determinants, namely modern contraception prevalence (mCP) and median age at first marriage, align with the changes observed in the spatial patterns of district-level fertility. Proximate determinants refer to the biological and behavioral factors through which distal determinants influence fertility. Additionally, since proximate determinants are directly linked to fertility, variations in one or more of these factors are inherently connected to differences and changes in fertility levels and trends over time (Bongaarts, 2015). This thesis demonstrates that both the accelerated increases in the two selected proximate determinants and the faster pace of fertility decline at the district level were particularly notable in the districts of the Amhara region in 2011 and 2016. Previous studies have indicated that effective family planning programs can enhance the utilization of modern contraceptives, which, in turn, can reduce unintended births and desired family size through informative messages about the advantages of family planning (Bongaarts, 2014, Casterline and Sinding, 2000, Cleland et al., 2014). Moreover, the rate of fertility decline can be accelerated through the diffusion of modern family planning information among populations sharing a common language or culture, regardless of their socioeconomic conditions (Bongaarts and Watkins, 1996).

The analysis conducted in Chapter 4 reveals that the geography of ethnolinguistic diversity remained relatively unchanged between 2000 and 2016, with consistently low levels of ethnolinguistic diversity observed in the districts of the Amhara region. It is noteworthy that the Amhara people, the second-largest ethnolinguistic group in Ethiopia with an estimated population of 20 million in 2020, constitutes approximately 27% of the total national population. Additionally, about 93% of the population in the Amhara region identifies as Ethiopian Orthodox Christians. This thesis demonstrates that two key proximate determinants, namely modern contraception prevalence (mCP) and age at first marriage, exhibited significant increases in the Amhara region, despite minimal changes in the proportions of women living in urban areas and with secondary education in the region. Several studies have reported the effectiveness of national family planning programs implemented in the Amhara region since 2004 (Olson and Piller, 2013, Tegegne et al., 2020, Halperin, 2014). Furthermore, in Ethiopia, where the traditional association between marriage and childbearing remains strong compared to many other Sub-Saharan African countries, delayed marriage is a prominent driver of fertility decline (Rogers and Stephenson, 2018). Median age at first marriage in the Amhara region consistently ranked the lowest across the eleven regions in Ethiopia, with values of 14.3 years in 2000 and 15.7 years in 2016 (appendix 4). Additionally, within the Orthodox culture prevalent in the Amhara region, extramarital childbearing is uncommon, and virginity is a religious prerequisite for marriage. Hence, the cultural and religious link between marriage and childbearing, coupled with the effective implementation of family planning programs and the shared language (Amharic) in the region, likely contributed to the accelerated fertility decline observed in the districts of the Amhara region in 2011 and 2016.

Overall, the rapid adoption of modern contraceptive methods and the delay in marriage observed in the Amhara region in 2011 and 2016 can be attributed to the effective implementation of family planning programs and the spatial diffusion effects facilitated by common ethnolinguistic contexts. These factors could have directly accelerated the pace of fertility decline in numerous districts of the Amhara region, irrespective of their socioeconomic conditions. This finding underscores the importance of considering both proximate and distal determinants of fertility when exploring geographical variations in district-level fertility in Ethiopia. It supports the notion that changes in reproductive behaviour, such as delayed marriage and increased contraceptive use, can be reinforced by diffusion effects through cultural and linguistic similarities, regardless of socioeconomic conditions.

3) Role of geographical location and distance on geographical variations in district-level fertility

This thesis highlights the role of geographical location of and distance between districts in describing geographical variations in district-level fertility in Ethiopia. The findings of this study revealed that even after accounting for the key selected proximate and distal determinants, spatial autocorrelation of Total Fertility Rate (TFR) at the district level became more pronounced in recent years. In simpler terms, districts that are closer to each other tend to exhibit more similar fertility levels compared to districts that are farther apart.

The diffusionist approach argues that the spread of information about birth control through interpersonal communication can be influenced by both geographical and cultural proximity (Bongaarts and Watkins, 1996). This implies that the recent variations in district-level fertility in Ethiopia by 2011 are not solely associated with differences in the key selected proximate and distal determinants, but also influenced by the geographical location and distance between districts. Interpersonal communication with individuals living nearby becomes particularly important when alternative communication channels, such as mobile phones, radio, and TV, are not readily available. According to the 2016 Ethiopia Demographic and Health Survey (DHS), only 27% of women aged 15-49 owned mobile phones, and approximately three out of four women (74%) had limited access to mass media, including radio, TV, and newspapers on a weekly basis. Consequently, in many Ethiopian societies, local communication serves as a crucial channel, making the location and distance between districts significant factors in describing geographical variations. In other words, the fertility level in a district may be influenced by the characteristics of neighbouring districts in addition to its own characteristics. Therefore, regression models that overlook the role of spatial location and distance can be biased.

If the spatial location by itself matters in geographical variations in district-level fertility, there are possibilities that the process of fertility decline can vary from district to district. Therefore, I examined whether the effects of the key selected determinants on fertility are homogeneous or heterogeneous across districts by using the local regression analysis (GWR) that uses distance-based weighting. The findings indicated that the relationships between district-level TFRs and the two selected proximate and two socioeconomic factors generally exhibited negative associations across most districts between 2000 and 2016, albeit with varying magnitudes. However, the GWR model revealed spatial heterogeneity in the relationship between TFRs and ethnolinguistic diversity, both in terms of magnitude and direction, across districts. Positive relationships were observed in the Amhara, Benishangul-Gumuz, and Gambela regions, while negative relationships were found in the three urban regions and the Afar and Somali regions. From a diffusionist perspective, the positive relationship aligns with expectations, as it suggests that ethnolinguistic similarity can reinforce the pace of fertility decline.

Therefore, considering that the geography of ethnolinguistic diversity in Ethiopia remained similar between 2000 and 2016, the presence of ethnolinguistically homogeneous contexts in the Amhara region (characterised by a low diversity index) may facilitate fertility decline (resulting in low fertility rates). Conversely, ethnolinguistically diverse contexts in the Benishangul-Gumuz and Gambela regions (exhibiting a high diversity index) are associated with higher fertility levels (resulting in high fertility rates).

On the other hand, the negative relationships observed in certain regions could be attributed to either higher fertility levels in ethnolinguistically homogeneous regions or lower fertility levels in ethnolinguistically diverse regions. For example, the Somali and Afar regions, despite being ethnolinguistically homogeneous, consistently exhibited higher fertility levels compared to other regions. The mean ideal numbers of children were also notably high in these regions, with 5.6 for Afar and 10.6 for Somali regions, surpassing the national average of 4.5 in 2016 (ICF, 2016). Conversely, the consistently lower fertility levels observed in multi-ethnic urban areas could be attributed to the adaptation to urban lifestyles by both urban residents and rural-to-urban migrants. Unlike rural areas, urban areas provide access to multiple communication channels. Take Addis Ababa, the capital and a multi-ethnic city, as an example. A significant portion of the population in Addis Ababa can speak the official language, Amharic, and there are various communication channels available for exchanging information and ideas. The 2016 Ethiopia DHS indicates that 87% of women in Addis Ababa owned mobile phones, and only 14% had no access to mass media. Consequently, the lower fertility levels observed in urban areas can be attributed to both strong adaptation effects on urban residents and rural-to-urban migrants and diffusion effects facilitated by multiple communication channels that surpass the limitations of interpersonal communication constrained by geographical distance. This finding in urban areas supports the view that the degree to which geographical distance can foster fertility decline are negatively correlated with socioeconomic or urban status (Klüsener et al., 2019, Costa et al., 2021).

Previous research in HMICs has also investigated the role of geographical distance as an important moderator in fertility decline. These studies have demonstrated that the fertility level in a given place is associated with the fertility level in neighbouring areas (Montgomery and Casterline, 1993, Goldstein and Klüsener, 2014, Klüsener et al., 2019). However, spatial analysis of fertility change at the district level is still rare in sub-Saharan African (SSA) countries, where significant variations in socioeconomic and cultural contexts exist. This implies that fertility in a district can be influenced not only by its own characteristics but also by the geographical location and characteristics of nearby districts.

6.3. Strengths

In this section, I will discuss the strengths of the thesis. In the section 2.7, I summarised three research gaps in the study of the geographical variation in fertility in Ethiopia. The major strength of this thesis is to address those research gaps introduced in chapter 2:

1. Although districts (Admin 3) are essential administrative units for health policy implementation and delivery in Ethiopia, very little is known about geographical variations in fertility at the **district level** due to the shortage of district-level data in Ethiopia.

2. Although theoretical and empirical studies support the presence of spatial dependency and heterogeneity of fertility in most populations, fertility studies in SSA countries, including Ethiopia, often neglect the **spatial effects** on sub-national variations in fertility.

3. Recent spatial analysis of fertility in HMICs have revealed the spatially varying relationship between distal determinants and fertility levels at small-scale spatial units. However, these studies often exclude the **role of proximate determinants**, which are crucial factors accounting for geographical variations in fertility in SSA countries.

6.3.1. Use of freely available geo-referenced data to explore geographical variations in district-level fertility in Ethiopia

First, I used the Demographic and Health Survey (DHS) data, which can be accessed freely upon approval of a data request from the DHS website (https://dhsprogram.com/data/) , to describe demographic and health outcomes at the district level in Ethiopia. While the national health strategy in Ethiopia acknowledges the existence of geographical variations in fertility and health outcomes at the district level, there is a shortage of district-level data in the country to comprehensively depict the spatial distribution of district-level fertility. Census data can potentially offer the necessary demographic information at the district level. However, in Ethiopia, census data are collected approximately every 10 years or even longer intervals. The last census was conducted in 2007, and the fourth census, initially scheduled for 2017, has been officially postponed four times as of October 2022 due to political, financial, and COVID-19-related challenges. Furthermore, census data often focus solely on demographic indicators and do not encompass health indicators.

In order to overcome the lack of district-level data in Ethiopia, this study used data from the Ethiopia Demographic and Health Survey (EDHS). This analytical approach can be readily expanded and applied to future EDHS surveys as well as other countries surveyed by the DHS programme. While previous studies have utilised the EDHS to investigate geographical variations in Ethiopia, they have primarily focused on variations between the eleven regions, rural-urban areas, or presented high-resolution maps of health outcomes at the administrative level 2. To my knowledge, this is the first study to go beyond previous studies on geographical variations in fertility in Ethiopia by describing and exploring the variations at the level of the 981 districts.

Examining demographic and health outcomes at the district level is crucial in sub-Saharan Africa (SSA), including Ethiopia, as many SSA countries have decentralized healthcare services to the district government level. The Woreda-based Health Sector Plan (WBSHP) Ethiopia's national health policy explicitly emphasises this decentralization. While the estimates for the 981 woredas based on approximately 500-600 sets of fertility data from the DHS may not be highly precise, this thesis offers valuable insights into geographical variations in fertility at the district level in SSA. It also sheds light on the emergence of health inequalities in high burden settings at the district level in SSA. By focusing on the district level, this study contributes to a better understanding of the dynamics of health outcomes and highlights the importance of addressing health disparities within specific local contexts.

6.3.2. Combination of spatial interpolation and spatial regression methods for spatial effects

In this thesis, I have employed various spatial models to investigate the spatial effects on geographical variations in district-level fertility in Ethiopia between 2000 and 2016. The inclusion of spatial autocorrelation and heterogeneity is crucial in comprehending the geographical patterns of fertility within a country, as supported by both theoretical arguments and empirical evidence (Wang and Chi, 2017, Vitali and Billari, 2017, de Castro, 2007). Previous studies conducted in high- and middle-income countries have explored spatial autocorrelation and heterogeneity in local-level fertility using spatial regression models (Haque et al., 2019, Burillo et al., 2020, Obradovic and Vojkovic, 2021).

However, as discussed earlier, the lack of local-level data in many SSA countries hinders the direct application of traditional spatial regression models. In this study, a novel approach was employed to overcome this limitation. The study combined model-based geostatistics (MBG) with spatial regression models to explore two important spatial effects: spatial autocorrelation of fertility and spatially heterogeneous relationships. Specifically, a Bayesian MBG approach was utilised and implemented through a stochastic partial differential equations (SPDE) approach with integrated nested Laplace approximations (INLA). This approach, which has been recently used in DHS spatial reports, allowed for the prediction of district-level total fertility rate (TFR) and key related determinants for the 981 districts in Ethiopia between 2000 and 2016. The study then compared non-spatial and spatial linear regression models to examine the presence of spatial autocorrelation in district-level fertility, even after controlling for selected proximate and distal determinants. Additionally, a geographically weighted regression (GWR) model was employed to uncover the spatially heterogeneous relationships between district-level TFR and the proximate and distal determinants.

By combining these different spatial methods, the study was able to investigate how geographical variations in district-level fertility are influenced not only by the socioeconomic and cultural characteristics of each district but also by the spatial autocorrelation and heterogeneity of fertility in relation to its proximate and distal determinants.

6.3.3. Spatially varying relationship between district-level fertility and both proximate and distal determinants of fertility

This study aimed to investigate the association between both proximate and distal determinants and the geographical variations in district-level fertility, as well as to examine whether these relationships exhibit spatial heterogeneity. While previous studies in high- and middle-income countries have utilised spatial models to explore the spatially heterogeneous relationships between district-level fertility and distal determinants, they have often overlooked the role of proximate determinants. However, in the context of sub-Saharan Africa (SSA), proximate determinants remain crucial in understanding geographical variations in fertility, as distal determinants can only impact fertility levels through their influence on proximate determinants. Moreover, measures such as modern contraceptive prevalence (mCP) and age at first birth still exhibit low levels and geographic variations in many SSA countries (Finlay et al., 2018, Laelago et al., 2019). To address these gaps, this study examined both key selected proximate and distal determinants to explore the potential spatial variability in the relationships between fertility and these determinants across districts.

Methodologically, the geographically weighted regression (GWR) model was employed as the primary methodology. The GWR model allows for the estimation of a local regression equation for each district in the dataset, thereby evaluating the relationships between variables at the district level. This approach involves incorporating the outcome variable (TFR) and explanatory variables from neighbouring districts that are in proximity to the target district. A neighbourhood (also known as a bandwidth) is the distance band (fixed) or number of neighbours (adaptive) used for each local regression equation. I adopted the adaptive method to determine bandwidth size, as the distance between districts in Ethiopia is not equal. This ensures that the GWR model captures the spatially heterogeneous relationships between fertility and its influencing factors across districts. By conducting a local analysis, the study enables the interpretation of these spatially heterogeneous relationships at the district level, providing insights from different perspectives of demographic theories.

6.4. Limitations

Limitations have been described at the end of each chapter. In this section, limitations that should be considered for the overall conclusions of this thesis will be presented and discussed.

6.4.1. Limited set of study variables

This study has limitations in terms of the selected proximate and distal determinants used to explore geographical variations in district-level fertility in Ethiopia. The Bongaarts model of proximate determinants of fertility identified four factors - contraception, marriage/cohabitation, induced abortion, and postpartum infecundability - to quantify their impact on fertility (Bongaarts, 2015). However, this study focused only on two proximate determinants, contraception and marriage exposure, excluding abortion and postpartum infecundability, because previous research has indicated that the differences in inhibiting effects of contraception and marriage on fertility between regional states in Ethiopia are larger compared to the other two determinants (Laelago et al., 2019, Alazbih et al., 2017). Nevertheless, it is important to acknowledge that postpartum infecundability has been considered an additional contributing factor to the recent fertility decline in rural Ethiopia (Lailulo and Sathiya Susuman, 2018, Todd and Lerch, 2021). Moreover, although statistics on abortion are frequently unavailable from the Ethiopia Demographic and Health Surveys due to its legal status in many low-income countries (Westoff, 2008), studies in six Asian countries have shown that induced abortion has a significant impact on fertility reduction among poor women (Majumder and Ram, 2015). Despite the increasing number of legal abortions in sub-Saharan Africa, a substantial number of abortions still occur illegally outside of health facilities (Moore et al., 2016, Singh et al., 2010). Therefore, the inclusion of these two excluded proximate determinants could provide additional insights into explaining geographical variations in district-level fertility in Ethiopia.

Furthermore, while other studies investigating geographical variations in fertility have examined a broader range of distal determinants, such as various economic indicators, mobile phone ownership, and media exposure, this thesis focused on only three distal determinants. This decision was based on the fact that, according to the latest reports from the Demographic and Health Surveys, Ethiopia had the largest differentials in fertility based on residential areas (rural and urban) and female educational levels among all the countries surveyed by the DHS programme (see Appendix 3). However, it should be noted that the aim of this DrPH thesis is not to identify the determinants of fertility. Instead, the focus is on describing the patterns of selected key determinants and gaining insights into their influence on the geographical variations in district-level fertility in Ethiopia.

6.4.2. Simple mean zonal statistics for aggregation of point estimate interpolated surface to districts

One primary method of utilising modelled surfaces for decision-making is through the aggregation of point estimate model surface pixels to relevant administrative units or policy areas (Janocha et al., 2021). In this study, a new map was created at the district level, illustrating the mean estimate value for each district. The aggregation process involved averaging the point estimates, which were aggregated to the 981 districts using simple mean zonal statistics. Recent research has also employed population-weighted mean zonal statistics, which use similar techniques but account for the estimated population in each grid square(Burgert-Brucker et al., 2016b). These studies often utilized population counts raster data from sources like the WorldPop to apply population-weighted aggregation when generating estimates at the administrative level in Ethiopia (Mayala et al., 2019b, Amoah et al., 2021). However, it is crucial to use the correct reference population layer, considering age-group and gender, in order to accurately estimate the denominator.

In this study, the denominator for estimating the fertility rates would ideally be the population counts of women aged 15-49 years at small grid squares in Ethiopia between 2000 and 2016. However, population data specifically for the female age group during that time period are not available. Given the relatively small size of districts in Ethiopia, it was assumed that there is an even distribution of populations within each district. Therefore, the simple mean zonal statistics were used for the aggregation. However, it should be noted that some districts in Ethiopia may have an uneven distribution of populations within their administrative units. The use of population-weighted mean zonal statistics, by utilizing the correct reference population layer, could potentially improve the estimation at the district level.

6.4.3. The absence of covariates in model-based geostatistics

It is acknowledged that including more covariates in the analysis can potentially improve the predictive ability of model-based geostatistics (Utazi et al., 2021). However, this study did not incorporate geospatial covariates in the model-based geostatistical approach. As discussed in Chapter 3, the goal of generating maps in this study was to produce ‘standardised maps’ for Ethiopia rather than aiming for the "best possible map." This standardised approach allows for comparisons between maps from different countries or different time periods, even when the availability of covariate datasets varies. Additionally, excluding geospatial covariates was intentional to focus on exploring associations between outcome and explanatory variables. Moreover, several studies in the field of global health utilised geospatial covariates obtained from publicly available remote sensing source, including land surface temperature, enhanced vegetation index, average monthly rainfall so on, for disease prevalence or environmental mapping (Giorgi et al., 2021, Huang et al., 2017, Reiner Jr et al., 2020). These covariates can primarily explain variation in data for certain diseases like malaria, but their contribution to explaining complex or socially driven health outcomes like HIV and fertility is relatively limited (Mayala et al., 2020). Therefore, geospatial covariates were not used in this study. However, it is recognised that incorporating appropriate geospatial covariates could potentially improve the accuracy of standardised maps.

6.4.4. Uncertainty

The results in Chapter 5 need to be interpreted carefully since they were derived from MBG-based estimates of outcome and explanatory variables at the district level in Chapter 4, rather than from raw data of EDHS. The stability and precision of the MBG-based estimates in Chapter 4 are affected by two elements of uncertainty. The first uncertainty is related to the sampling issues of the EDHS, and the second uncertainty is inherent in MBG.

Firstly, uncertainty can arise from displaced EDHS cluster locations and small sample sizes at each primary sampling unit (PSU). The locations of PSUs used in the spatial modelling process represent estimated centres of clusters of households. This means that the actual household locations are unknown, and the point locations for PSUs represent areas of varying size. In urban areas, these point locations are geo-masked within a range of 0-2 km, and in rural areas, the range is 0-5 km, with 1% of rural locations extending up to 10 km. Therefore, the displaced EDHS cluster locations can lead to positional uncertainty in accurately estimating values at a certain location. In addition, small sample sizes, consisting of approximately 25-35 women, were collected from each PSU. As a result, estimates of study variables at PSUs may be subject to sampling errors, as smaller samples are less likely to represent the population. Previous studies have shown that MBG can moderate the impact of small samples and DHS point displacement (Gething et al., 2015, Mayala et al., 2020). This is because MBG enables the borrowing of information from neighbouring areas and incorporates covariate information, resulting in the smoothing or shrinking of extreme or misleading values. However, it should be acknowledged that the DHS sampling itself is subject to uncertainty in the estimates at the PSU level. In particular, recent studies have developed advanced geostatistical models that account for the positional uncertainty of DHS data to improve predictive measures (Fonterre et al., 2018, Wilson and Wakefield, 2021). Therefore, such geostatistical approaches should be considered in future studies.

Secondly, an important element of the MBG-based surface output is the uncertainty estimates. I presented uncertainty maps with Bayesians credible intervals for each predicted pixel value. (Figure 4.5) (Janocha et al., 2021). The map revealed that the level of uncertainty tends to be relatively higher in the Afar and Somali regional states, as well as the southern parts of Ethiopia. Bayesians credible intervals of uncertainty that grow where there is heterogeneity in data (e.g., wide range of total fertility rate) or when there are no data to support prediction (Mayala et al., 2020). Therefore, the higher uncertainty can be attributed to a smaller number of PSUs in those areas and the concentration of excluded PSUs without GPS coordinate information in the Afar and Somali regions in 2011 and 2016 (Appendix 2). In addition, the uncertainty map showed that uncertainty at unsampled locations for TFR and mCP increased from 2000 to 2016, which may be due to the widened geographical variations in TFR and mCP in recent years. Therefore, although the MBG-based estimates from this study provide valuable information on geographical variations in district-level fertility, it is important evaluate the stability and precision of estimates at specific locations on the map (Figure 4.4) by examining the uncertainty surface (Figure 4.5). Higher uncertainty in a location indicates less accuracy in the model's estimation of the indicator value, while lower uncertainty suggests a better estimate for that location. Understanding and acknowledging this uncertainty is crucial as it provides information about the model's reliability and the variations in uncertainty across the country. Hence, when individuals or policy makers want to use the result in Chapter 4 and 5, they should determine an acceptable level of uncertainty based on the context, and their tolerance for uncertainty.

Despite the uncertainty, I also found that the aggregated values at the regional level generally fall within the 95% confidence interval of the estimates derived directly from the EDHS survey data files (Figure 2.3 and Figure 4.2). Hence, it could be expected that aggregated values at the district level may fall within the 95% confidence interval of the true values of district-level TFR, which should be further validated with available district-level data. Ideally, to obtain more precise estimates at the district level, larger and more comprehensive datasets would be needed, which can be costly and time-consuming. Alternatively, a potential solution could involve combining multiple data sources, including the Ethiopia DHS, Census, and Performance Monitoring for Action (PMA) data, and incorporating them within a Bayesian model-based geostatistical (MBG) framework to estimate fertility and its determinants at the district level.

6.4.5. Lack of temporal analysis

This study is conducted as a cross-sectional analysis, and therefore, all the analysis methods were implemented separately for each EDHS survey. Consequently, the findings from this study cannot be used to draw conclusions about the role of spatial patterns and distributions in driving district-level fertility changes over time. It is important to recognise that fertility levels not only vary across different geographical locations but also differ across different sampling periods. The relationships between district-level fertility and the influencing proximate and distal determinants may be influenced by both spatial and temporal factors. Hence, studies investigating geographical variations in fertility may require both spatial and temporal perspectives. One advantage of using a model-based geostatistical approach is its flexibility in analysing the spatio-temporal dynamics of demographic phenomena. Additionally, the geographically weighted regression (GWR) model has been further extended to incorporate temporal information into the spatial heterogeneity of health outcomes through the geographically and temporally weighted regression (GTWR) model(Chen et al., 2021, Guo et al., 2021).

However, it is important to note that this DrPH thesis focuses on investigating geographical variations in fertility rather than examining the temporal trends of district-level fertility. Nevertheless, this study utilised four rounds of the EDHS data and presented the spatial patterns and relationships of district-level fertility in relation to proximate and distal determinants for the years 2000, 2005, 2011, and 2016. It can be reasonably assumed that the cross-sectional effect of spatial correlation on geographical variations in fertility may reflect the influence of spatial correlation on fertility from previous years.

6.5. Implications

6.5.1. Increasing importance of district-level analysis in geographical variations in fertility

This thesis reinforces the significance of district-level analysis in describing geographical variations in fertility in Ethiopia in recent years. The presence of such variations is an important characteristic of the country, as evidenced by substantial differences in fertility rates between urban and rural areas as well as regional states. While previous studies on geographical variations in fertility in Ethiopia have often focused on differences between urban and rural areas or regional states, this study explores variations at the district level.

The first question addressed in this thesis is whether there are geographical variations in fertility at the district level between 2000 and 2016. Figure 4.4 clearly demonstrates the emergence of geographical variations in district-level fertility in recent years. Studies conducted in the early 2000s without district-level analysis may have provided a relatively adequate description of geographical variations in fertility in Ethiopia at that time. However, without district-level analysis in more recent years, such studies would be unable to capture the emerging variations in fertility at the district level.

From a national health policy perspective, the Health Sectoral Transformation Plan II (HSTP II) (2020/21 – 2024/25) in Ethiopia explicitly recognises the high level of variations in health and demographic outcomes between districts. The HSTP II aims to achieve universal coverage and equity of essential health services across districts in the country. To address the delivery of health services, Ethiopia has implemented a decentralised health system, where the district (woreda) serves as a basic administrative unit with the authority to manage budgets and allocate resources according to their needs and strategic plans aligned with the national HSTP II. The transformation of each woreda is expected to establish an accountable and transparent health system that fosters meaningful community participation and data-driven decision-making.

Despite the importance of a decentralised health system through the woreda-based transformation in Ethiopia, challenges remain regarding the availability of complete and timely data for evidence-based decision-making at the district level. While the population and housing census provides district-level data, the last census in Ethiopia was conducted in 2007, and the next census is still in preparation as of July 2022. Moreover, census data does not provide information on health outcomes. The health and demographic surveillance system (HDSS) offers demographic and health data for only seven districts, which is insufficient to describe geographical variations in fertility across the entire country. Although recent DHS spatial reports provide health outcomes at the Admin 2 level, I argue that the woreda level is the more essential administrative unit.

This study reveals that while regional and urban-rural differences are important aspects of geographical variations in fertility, there are also substantial variations between districts within the same regions, particularly in recent years. This implies that now is the time provide more district-level data to bridge the gaps between policy priorities and evidence, thereby enhancing decentralised health services in Ethiopia. Otherwise, it could hinder efforts to understand the factors determining fertility in the country and impede the planning and monitoring of decentralized health programs at the district level.

6.5.2. Spatial effect on geographical variations in district level-fertility

This thesis underscores the increasing significance of spatial autocorrelation in fertility as a key aspect of geographical variations in Ethiopia, particularly in recent years. It reveals that socioeconomic and cultural characteristics vary among districts, even within the same region, and that fertility in a district is influenced by both its own characteristics and those of neighbouring districts. In Chapter 2, two possible explanations for geographical variations in fertility were introduced: the adaptationist and diffusionist approaches. The adaptationist approach suggests that fertility is influenced by socioeconomic factors within specific areas. On the other hand, the diffusionist approach posits that fertility intentions are influenced by the spread of ideas and practices, especially among individuals who share similar ethnolinguistic backgrounds. In the case of Ethiopia, where the geographical boundaries of administrative units largely align with the distribution of over 90 different ethnolinguistic groups, the diffusionist perspective becomes particularly relevant. Therefore, considering the spatial dimension becomes crucial for understanding and analysing geographical variations in fertility in Ethiopia.

As explained in Chapter 2, ethno-linguistic identity is the primary criterion shaping the administrative boundaries of the country (Admin 1). The Ethiopian constitution explicitly states that every ethno-linguistic group has the right to establish self-administrative areas starting at the district (woreda) level, and at the zonal and regional levels depending on the size of each ethno-linguistic group (Ethiopia, 1995). Recently, there have been changes in the ethnolinguistic geography of Ethiopia. New ethnolinguistic groups have established their own regions since 2016. Between 1994 and 2020, there were only eleven regions (nine regions and two chartered cities) and one official language (Amharic). However, this has recently begun to change.

Firstly, the Sidama and Southwest regions were newly established in June 2020 and November 2021, respectively, following the successful votes by the Sidama ethnolinguistic group and the southwest Ethiopian people to create their own regional states from the Southern Nations, Nationalities, and Peoples' (SNNP) region. Additionally, two more regions (Damotic & Omotic region and Northern & central region) are ready to become independent from the SNNP region and establish new regions. Among these newly established or establishing four regions, the Sidama region is ethnolinguistically homogeneous, but the other three regions will be multi-ethnic regions. Consequently, the SNNP region will eventually be divided into four regions, resulting in a total of twelve regions and two chartered cities in Ethiopia. Secondly, in 2021, the Ethiopian government added four additional official languages (Tigray, Oromo, Hara, Somali) to the initial official language (Amharic). Therefore, Ethiopia currently has five official languages.

These recent changes in geography and language in Ethiopia are likely to influence spatial and communication interactions, which can in turn affect reproductive health behaviours and fertility intentions according to the diffusionist approach. These changes suggest that ethnolinguistic-based geography could play a greater role in Ethiopia, highlighting the need for increased attention to the spatial aspects of district-level fertility in the country.

6.6. Future work

There are many ways to extend the work presented in this thesis to better understand geographical variations in fertility at the district level in SSA.

6.6.1. Onset and pace of fertility decline at district level in SSA using spatiotemporal model.

The spatial method used in this thesis provides an approach to describe changes in the spatial patterns of district-level fertility in Ethiopia between 2000 and 2016. However, employing spatio-temporal methods would offer additional benefits as they allow for the examination of persistent patterns and identification of atypical patterns across both space and time. By utilising spatio-temporal methods, we can not only a) predict district-level fertility but also b) explore the spatially and temporally varying relationships between district-level fertility and factors influencing fertility in Ethiopia.

Firstly, the spatial modelling of geostatistical data used to predict fertility rates for 981 districts in this study can be extended to incorporate the time dimension. By including time into the model, spatio-temporal modelling can account for variations in the fertility trend across different districts over time. Spatio-temporal mapping models are commonly applied in disease surveillance studies, where the focus is on predicting disease occurrence in observed and unobserved areas over time. In fertility studies, demographers have long studied the onset and pace of fertility decline in Sub-Saharan Africa (SSA). Several studies have revealed that fertility decline in Africa has been delayed compared to other regions, and the pace of decline has been relatively slow. However, these studies have primarily focused on fertility trends at the national level. It would be particularly interesting to examine how the onset and pace of fertility decline differ between districts within SSA countries. By analysing fertility patterns at the district level, we could gain a more detailed understanding of the variations and nuances in fertility behaviours and outcomes.

Secondly, the geographically weighted regression (GWR) model used to explore spatially varying relationships between district-level fertility and determinants in this thesis can be expanded to incorporate both spatial and temporal weights through geographically and temporally weighted regression (GTWR). As mentioned earlier, the onset and pace of fertility decline may differ at the district level, making it important to investigate the variations in district-level fertility in relation to proximate and distal determinants over both space and time. While this thesis demonstrates the spatially varying relationship between fertility and variables, their relationship can also vary over time. Incorporating both time and space will enhance our understanding of how the onset and pace of fertility decline at the district level vary spatially and temporally in Ethiopia.

6.6.2. Application of district-level analysis to further DHS data and other countries

This study used freely available DHS data and focused solely on Ethiopia during four specific years (2000, 2005, 2011, and 2016). A similar study could be extended in two ways: 1) by using the upcoming EDHS to explore fertility patterns within Ethiopia, and 2) by expanding the research to other countries in Sub-Saharan Africa (SSA) to generalise the geographical variations in district-level fertility across the region.

Firstly, the study could be extended within Ethiopia using the subsequent EDHS surveys. Since the 2016 EDHS, the interim 2019 EDHS data has become available, and the data collection for the 2022 EDHS is currently underway. It is worth noting that Ethiopia has experienced significant events between 2019 and 2022, including severe civil unrest and the COVID-19 pandemic, which have greatly impacted reproductive health services in the country. The political landscape in Ethiopia underwent a change in 2019, with the dissolution of the Ethiopia People's Revolutionary Democratic Front (EPRDF) and the formation of the Prosperity Party (PP). The Ethiopia People's Revolutionary Democratic Front (EPRDF), which dominated Ethiopia politics from 1991 to 2019 by the political coalition of four ethnic- and territorial- based parties (Oromo Democratic Party (ODP), Amhara Democratic Party (ADP), Southern Ethiopian People's Democratic Movement (SEPDM) and Tigray People's Liberation Front (TPLF)) was dissolved in 2019 and the new political party, the Prosperity Party (PP), was formed in December 2019. In addition to the previous political coalition, five additional ethnic-territorial based parties (The Harari National League (HNL), the Benishangul-Gumuz People's Democratic Unity Front (BGPDUF), the Afar National Democratic Party (ANDP), the Gambela People's Democratic Movement (GPDM) and the Ethiopian Somali People's Democratic Party (ESPDP)) joined the PP. However, the TPLF, which led the EPRDF between 1991-2019, disagreed on the formation of the new party and the tension between the Prosperity Party and the TPLF eventually triggered the current northern Ethiopia Crisis since November 2020. The northern Ethiopia crisis has had a devastating effect on healthcare delivery, resulting in the displacement of millions of people and the collapse of the health information and service delivery systems in the affected regions. Additionally, the COVID-19 pandemic, although relatively less severe in SSA compared to other regions, has still caused disruptions to reproductive health services due to mobility restrictions and economic insecurity. Therefore, further studies should be conducted to monitor and assess the dual impacts of the northern Ethiopia crisis and COVID-19 on reproductive health outcomes in specific districts in the Amhara and Tigray regions.

Secondly, this thesis demonstrates the existence of spatial autocorrelation and changes in district-level fertility in Ethiopia. To further explore the spatial correlation of district-level fertility, the spatial analysis using DHS data can be extended to include neighboring districts in Ethiopia's neighbouring countries. For instance, Ethiopia shares a border with Kenya, and the Kenya Demographic and Health Surveys (DHS) have been conducted in various years, including 1989, 1993, 1998, 2003, 2008-09, 2015, and the ongoing 2022 survey. Previous research has shown that neighbouring districts from different countries, particularly at national borders, often exhibit similar fertility behaviours due to shared culture, language, and history, regardless of their socioeconomic characteristics (Das et al., 2020, Campisi et al., 2020). These findings suggest the possibility of cross-border diffusion of fertility practices between countries. Therefore, it would be interesting to investigate whether a spatial relationship of district-level fertility can be observed across the Ethiopian-Kenyan border using data from both the Ethiopia and Kenya DHS surveys. Particularly, the Moyale district is located between the border of Ethiopia and Kenya. In addition, the Moyale district is also located at the border between Oromo and Somali regions in Ethiopia. Thus, the name of 'Moyale' appears in a map of three regions. One (Moyale woreda) in the Dhawa Zone in Somali region and another one (Moyale woreda) in the Borena zone in Oromia region of Ethiopia, and another one (Moyale town) located in the Eastern province in Kenya. Therefore, it would be interesting to describe and explore the spatial aspects of fertility in these districts, which may not be fully captured by national or regional-level studies.

Thirdly, district-level analysis can be extended to countries that have implemented effective national family planning (FP) programs to investigate the association between FP programs and fertility at the district level. The desire for large families is commonly cited as a contributing factor to the high fertility rates in SSA (Bongaarts, 2020). However, investments in FP programs have often been insufficient in many SSA countries, as policymakers tend to perceive FP programs as having minimal impact in settings with limited demand for birth limitation (Janocha et al., 2021). Ethiopia, Rwanda, and Malawi are frequently mentioned as countries that have successfully implemented effective national family planning programs, leading to significant increases in contraceptive use (Cates Jr and Maggwa, 2014). The DHS has been conducted six times in Rwanda (1992, 2000, 2005, 2010, 2015, 2019) and five times in Malawi (1992, 2000, 2004, 2010, 2015). In terms of reproductive health outcomes, Ethiopia's fertility rate decreased from 5.7 in 2000 to 4.6 in 2016, accompanied by an increase in modern contraceptive prevalence from 6.3% in 2000 to 35.3% in 2016. Similarly, the total fertility rates (TFRs) of Rwanda and Malawi were 5.8 and 6.3 in 2000, and decreased to 4.2 and 4.4 in 2015, respectively. Moreover, the modern contraceptive prevalence increased from 5.7% and 26.1% in 2000 to 47.5% and 58.1% in 2015 for Rwanda and Malawi, respectively. Therefore, further studies can compare the spatial patterns of fertility and modern contraceptive prevalence at the district level among these three countries to investigate the effects of family planning programs on district-level fertility in association with socioeconomic and cultural contexts.

6.7. Concluding remarks

This DrPH thesis presents evidence of geographical variations in fertility at the district level in Ethiopia, and highlights the widening of these variations in fertility between 2000 and 2016. This DrPH reveals that district-level total fertility rates (TFRs) have increasingly diverged from their respective regional TFRs in recent years, and this divergence is associated with the spatial spread of fertility decline from

Addis Ababa to the northern and western parts of Ethiopia. The key driving factors behind this spatial spread are the modern contraceptive prevalence and age of first marriage, which are the two key proximate determinants of fertility, rather than the socioeconomic factors of urbanisation and female education. Furthermore, this thesis reveals the presence of spatial autocorrelation in district-level fertility and uncovers spatial heterogeneity in the relationship between district-level fertility and ethnolinguistic diversity within Ethiopia. The findings indicate that lower fertility levels are observed in multi-ethnic urban regions and the ethnolinguistically homogeneous Amhara region. These observations support both the diffusionist view, which suggests that the transmission of new ideas within common linguistic areas can accelerate fertility decline, and the adaptationist view, which suggests that the impact of geographical distance on fertility decline can be mitigated by socioeconomic or urban status. Notably, socioeconomic and cultural contexts vary significantly between districts in many countries in Sub-Saharan Africa (SSA). This thesis emphasises that fertility within a district can be influenced not only by its own characteristics but also by the geographical location of the district and the characteristics of nearby districts. Therefore, it is crucial to analyse fertility patterns at lower geographical levels in SSA, as higher-level analyses may overlook important spatial patterns that can inform public health intervention decisions. Overall, this DrPH thesis provides additional insights into how spatial factors, along with socioeconomic and cultural characteristics, and reproductive behaviours within districts, collectively shape geographical variations in fertility in Ethiopia.

6.8. References

ALAZBIH, N. M., TEWABE, G. N. & DEMISSIE, T. D. 2017. Contraception and fertility transition in AMHARA National Regional State of ETHIOPIA: an application of BONGAARTS'model. *Fertility research and practice,* 3**,** 12.

ALEM, A. Z., YESHAW, Y., KEBEDE, S. A., LIYEW, A. M., TESEMA, G. A., AGEGNEHU, C. D. & TESHALE, A. B. 2020. Spatial Distribution and Determinants of Early Marriage among Married Women in Ethiopia: A spatial and Multilevel Analysis. *BMC Women's Health,* 20**,** 1-13.

AMOAH, B., FRONTERRE, C., JOHNSON, O., DEJENE, M., SEIFE, F., NEGUSSU, N., BAKHTIARI, A., HARDING-ESCH, E. M., GIORGI, E. & SOLOMON, A. W. 2021. Model-based geostatistics enables more precise estimates of neglected tropical-disease prevalence in elimination settings: mapping trachoma prevalence in Ethiopia. *International journal of epidemiology*.

BECKER, G. S. 1960a. An economic analysis of fertility. *Demographic and economic change in developed countries.* Columbia University Press.

BEHRMAN, J. A. 2015. Does schooling affect women's desired fertility? Evidence from Malawi, Uganda, and Ethiopia. *Demography,* 52**,** 787-809.

BENZA, M., WEEKS, J. R., STOW, D. A., LõPEZ-CARR, D. & CLARKE, K. C. 2017. Fertility and urban context: A case study from Ghana, West Africa, using remotely sensed imagery and GIS. Population, Space and Place, 23, e2062.

BONGAARTS, J. 2014. The impact of family planning programs on unmet need and demand for contraception. *Studies in family planning,* 45**,** 247-262.

BONGAARTS, J. 2015. Modeling the fertility impact of the proximate determinants: Time for a tune-up. *Demographic Research,* 33**,** 535-560.

BONGAARTS, J. 2020. Trends in fertility and fertility preferences in sub-Saharan Africa: the roles of education and family planning programs. *Genus,* 76**,** 1-15.

BONGAARTS, J. & BRUCE, J. 1995. The causes of unmet need for contraception and the social content of services. *Studies in family planning***,** 57-75.

BONGAARTS, J. & WATKINS, S. C. 1996. Social interactions and contemporary fertility transitions. *Population and development review***,** 639-682.

BRYANT, J. 2007. Theories of fertility decline and the evidence from development indicators. *Population and development review***,** 101-127.

BURGERT-BRUCKER, C. R., DONTAMSETTI, T., MASHALL, A. & GETHING, P. 2016b. Guidance for use of The DHS Program modeled map surfaces. *DHS Spatial Analysis Reports No. 14.* Rockville, Maryland, USA: ICF International.

BURILLO, P., SALVATI, L., MATTHEWS, S. A. & BENASSI, F. 2020. Local-Scale Fertility Variations in a Low-Fertility Country: Evidence from Spain (2002–2017). *Canadian Studies in Population,* 47**,** 279-295.

CAMPISI, N., KULU, H., MIKOLAI, J., KLþSENER, S. & MYRSKYLæ, M. 2020. Spatial variation in fertility across Europe: Patterns and determinants. *Population, Space and Place***,** e2308.

CASTERLINE, J. B. & SINDING, S. W. 2000. Unmet need for family planning in developing countries and implications for population policy. *Population and development review,* 26**,** 691-723.

CATES JR, W. & MAGGWA, B. 2014. Family planning since ICPD—How far have we progressed? *Contraception,* 90**,** S14-S21.

CHEN, Y., CHEN, M., HUANG, B., WU, C. & SHI, W. 2021. Modeling the spatiotemporal association between COVID-19 transmission and population mobility using geographically and temporally weighted regression. *GeoHealth,* 5**,** e2021GH000402.

CLELAND, J., HARBISON, S. & SHAH, I. H. 2014. Unmet need for contraception: issues and challenges. *Studies in family planning,* 45**,** 105-122.

COSTA, R., BOCQUIER, P. & EGGERICKX, T. 2021. The Contribution of Diffusion to the Fertility Transition in Belgium (1887–1934). *Population and Development Review*.
DAS, P., HUSAIN, Z. & GHOSH, S. 2020. Fertility behaviour in linguistic zones: revisiting the diffusion hypothesis in greater Bengal. *Asian Population Studies,* 16**,** 34-60.
DE CASTRO, M. C. 2007. Spatial demography: An opportunity to improve policy making at diverse decision levels. *Population research and policy review,* 26**,** 477-509.
EASTERLIN, R. A. 1975. An economic framework for fertility analysis. *Studies in family planning,* 6**,** 54-63.
EBRAHIM, O. A., ZELEKE, E. G. & MULUNEH, A. G. 2021. Geographic variation and associated factors of long-acting contraceptive use among reproductive-age women in Ethiopia: a multi-level and spatial analysis of Ethiopian Demographic and Health Survey 2016 data. *Reproductive health,* 18**,** 1-12.
ETHIOPIA 1995. *The 1995 Constitution of Ethiopia,* Federal Democratic Republic of Ethiopia, Federal Democratic Republic of Ethiopia.
EYASU, A. M. 2015. Multilevel modeling of determinants of fertility status of married women in Ethiopia. *American Journal of Theoretical and Applied Statistics,* 4**,** 19-25.
FINLAY, J. E., MEJïA-GUEVARA, I. & AKACHI, Y. 2018. Inequality in total fertility rates and the proximate determinants of fertility in 21 sub-Saharan African countries. *PloS one,* 13**,** e0203344.
FRONTERRE, C., GIORGI, E., & DIGGLE, P. (2018). Geostatistical inference in the presence of geomasking: a composite-likelihood approach. Spatial Statistics, 28:319–330.
GETHING, P., TATEM, A., BIRD, T. & BURGERT, C. 2015. Creating spatial interpolation surfaces with DHS data. *DHS Spatial Analysis Reports No. 11.* Rockville, Maryland, USA: ICF.
GIBSON, M. A. & GURMU, E. 2012. Rural to urban migration is an unforeseen impact of development intervention in Ethiopia. PLoS One, 7, e48708.
GIORGI, E., FRONTERRê, C., MACHARIA, P. M., ALEGANA, V. A., SNOW, R. W. & DIGGLE, P. J. 2021. Model building and assessment of the impact of covariates for disease prevalence mapping in low-resource settings: to explain and to predict. *Journal of the Royal Society Interface,* 18**,** 20210104.
GOLDSTEIN, J. R. & KLþSENER, S. 2014. Spatial analysis of the causes of fertility decline in Prussia. *Population and Development Review,* 40**,** 497-525.
GUO, B., WANG, Y., PEI, L., YU, Y., LIU, F., ZHANG, D., WANG, X., SU, Y., ZHANG, D. & ZHANG, B. 2021. Determining the effects of socioeconomic and environmental determinants on chronic obstructive pulmonary disease (COPD) mortality using geographically and temporally weighted regression model across Xi'an during 2014–2016. *Science of The Total Environment,* 756**,** 143869.
HALPERIN, D. T. 2014. Scaling up of family planning in low-income countries: lessons from Ethiopia. *The Lancet,* 383**,** 1264-1267.
HAQUE, I., DAS, D. & PATEL, P. P. 2019. Reading the geography of India's district-level fertility differentials: a spatial econometric approach. *Journal of Biosocial Science,* 51**,** 745-774.
HUANG, J., MALONE, B. P., MINASNY, B., MCBRATNEY, A. B. & TRIANTAFILIS, J. 2017. Evaluating a Bayesian modelling approach (INLA-SPDE) for environmental mapping. *Science of the Total Environment,* 609**,** 621-632.
ICF International. 2015. The DHS Program STATcompiler [Online]. USAID. Available: http://www.statcompiler.com [Accessed December 19 2021].
ICF International and Central Statistical Agency. Ethiopia DHS 2016. Addis Ababa, Ethiopia and Calverton, MD: Central Statistical Agency, ICF International, 2016.
JANOCHA, B., DONOHUE, R. E., FISH, T. D., MAYALA, B. K. & CROFT, T. N. 2021. Guidance and recommendations for the use of indicator estimates at subnational administrative level 2. *DHS Spatial Analysis Report 20.* Rockville, Maryland, USA: ICF.
KLþSENER, S., DRIBE, M. & SCALONE, F. 2019. Spatial and social distance at the onset of the fertility transition: Sweden, 1880–1900. *Demography,* 56**,** 169-199.
KULU, H. 2005. Migration and fertility: Competing hypotheses re-examined. *European Journal of Population/Revue européenne de Démographie,* 21**,** 51-87.

LAELAGO, T., HABTU, Y. & YOHANNES, S. 2019. Proximate determinants of fertility in Ethiopia; an application of revised Bongaarts model. *Reproductive health,* 16**,** 13.

LAILULO, Y. A. & SATHIYA SUSUMAN, A. 2018. Proximate determinants of fertility in Ethiopia: comparative analysis of the 2005 and 2011 DHS. *Journal of Asian and African Studies,* 53**,** 733-748.

MAJUMDER, N. & RAM, F. 2015. Explaining the role of proximate determinants on fertility decline among poor and non-poor in Asian countries. *PloS one,* 10**,** e0115441.

MAYALA, B. K., BHATT, S. & GETHING, P. 2020. Predicting HIV/AIDS at subnational levels using DHS covariates related to HIV. *DHS Spatial Analysis Reports No. 18.* Rockville, Maryland, USA: ICF.

MAYALA, B. K., DONTAMSETTI, T., FISH, T. D. & CROFT, T. 2019a. *Interpolation of DHS survey data at subnational administrative Level 2*, DHS Program, ICF.

MAYALA, B. K., DONTAMSETTI, T., FISH, T. D. & CROFT, T. N. 2019b. Interpolation of DHS survey data at subnational administrative Level 2. *DHS Spatial Analysis Reports No. 17.* Rockville, Maryland, USA: ICF.

MENGESHA, M. M., ASSEFA, N. & DESSIE, Y. Socio-demographic Determinants of Adolescent Fertility Vary in Urban and Rural Areas in Eastern Ethiopia: A Cross-sectional Study Based on Health and Demographic Surveillance System Data. 29th EPHA Annual Conference, 2018.

MONTGOMERY, M. R. & CASTERLINE, J. B. 1993. The diffusion of fertility control in Taiwan: Evidence from pooled cross-section time-series models. *Population studies,* 47**,** 457-479.

MOORE, A. M., GEBREHIWOT, Y., FETTERS, T., WADO, Y. D., BANKOLE, A., SINGH, S., GEBRESELASSIE, H. & GETACHEW, Y. 2016. The estimated incidence of induced abortion in Ethiopia, 2014: changes in the provision of services since 2008. *International perspectives on sexual and reproductive health,* 42**,** 111.

OBRADOVIC, S. L. & VOJKOVIC, G. 2021. The Impact of Spatial-Demographic Disparities on Fertility Variations in Serbian Municipalities—An Example of Hidden Content Research. *Hidden Geographies***,** 379.

OECD 2020. Rural Development Strategy Review of Ethiopia. *OECD*.

OLSON, D. J. & PILLER, A. 2013. Ethiopia: an emerging family planning success story. *Studies in Family Planning,* 44**,** 445-459.

PEZZULO, C., NILSEN, K., CARIOLI, A., TEJEDOR-GARAVITO, N., HANSPAL, S. E., HILBER, T., JAMES, W. H., RUKTANONCHAI, C. W., ALEGANA, V. & SORICHETTA, A. 2021. Geographical distribution of fertility rates in 70 low-income, lower-middle-income, and upper-middle-income countries, 2010–16: a subnational analysis of cross-sectional surveys. *The Lancet Global Health,* 9**,** e802-e812.

REINER JR, R. C., WIENS, K. E., DESHPANDE, A., BAUMANN, M. M., LINDSTEDT, P. A., BLACKER, B. F., TROEGER, C. E., EARL, L., MUNRO, S. B. & ABATE, D. 2020. Mapping geographical inequalities in childhood diarrhoeal morbidity and mortality in low-income and middle-income countries, 2000–17: analysis for the Global Burden of Disease Study 2017. *the lancet,* 395**,** 1779-1801.

ROGERS, E. & STEPHENSON, R. 2018. Examining temporal shifts in the proximate determinants of fertility in low-and middle-income countries. *Journal of biosocial science,* 50**,** 551-568.

SALVATI, L., BENASSI, F., MICCOLI, S., RABIEI-DASTJERDI, H. & MATTHEWS, S. A. 2020. Spatial variability of total fertility rate and crude birth rate in a low-fertility country: patterns and trends in regional and local scale heterogeneity across Italy, 2002–2018. *Applied Geography,* 124**,** 102321.

SINGH, S., FETTERS, T., GEBRESELASSIE, H., ABDELLA, A., GEBREHIWOT, Y., KUMBI, S. & AUDAM, S. 2010. The estimated incidence of induced abortion in Ethiopia, 2008. *International perspectives on sexual and reproductive health***,** 16-25.

SHAPIRO, D. & TENIKUE, M. 2017. Women's education, infant and child mortality, and fertility decline in urban and rural sub-Saharan Africa. *Demographic Research,* 37**,** 669-708.

TEGEGNE, T. K., CHOJENTA, C., FORDER, P. M., GETACHEW, T., SMITH, R. & LOXTON, D. 2020. Spatial variations and associated factors of modern contraceptive use in Ethiopia: a spatial and multilevel analysis. *BMJ open,* 10**,** e037532.
TESSEMA, Z. T., AZANAW, M. M., BUKAYAW, Y. A. & GELAYE, K. A. 2020. Geographical variation in determinants of high-risk fertility behavior among reproductive age women in Ethiopia using the 2016 demographic and health survey: a geographically weighted regression analysis. *Archives of Public Health,* 78**,** 1-12.
TESSEMA, Z. T. & TAMIRAT, K. S. 2020. Determinants of high-risk fertility behavior among reproductive-age women in Ethiopia using the recent Ethiopian Demographic Health Survey: a multilevel analysis. *Tropical Medicine and Health,* 48**,** 1-9.
TODD, N. & LERCH, M. 2021. Socioeconomic development predicts a weaker contraceptive effect of breastfeeding. *Proceedings of the National Academy of Sciences,* 118.
UTAZI, C. E., NILSEN, K., PANNELL, O., DOTSE-GBORGBORTSI, W. & TATEM, A. J. 2021. District-level estimation of vaccination coverage: Discrete vs continuous spatial models. *Statistics in Medicine,* 40**,** 2197-2211.
VITALI, A. & BILLARI, F. C. 2017. Changing determinants of low fertility and diffusion: A spatial analysis for Italy. *Population, Space and Place,* 23**,** e1998.
WANG, D. & CHI, G. 2017. Different Places, Different Stories: A Study of Spatial Heterogeneity of County-Level Fertility in China. *Demographic research,* 37**,** 493.
WATKINS, S. C. The fertility transition: Europe and the Third World compared. Sociological forum, 1987. Springer, 645-673.
WESTOFF, C. F. 2008. A new approach to estimating abortion rates. *DHS Analytical Studies No. 13.* Calverton, Maryland, USA: Macro International.

WILSON, K & WAKEFIELD, J. 2021. Estimation of health and demographic indicators with incomplete geographic information. *Spatial and Spatio-temporal Epidemiology*, 37:100421.

Integrating Statement

The DrPH at LSHTM aims to "equip its graduates with the knowledge and experience to deal with the particular challenges of understanding and adapting scientific knowledge in order to achieve public health gains as well as the analytical and practical skills required by managers and leaders in public health." I joined the DrPH programme in 2017 after having worked in the field of global health in low- and middle-income countries at various government and non-government organisations for a few years.

The DrPH at LSHTM programme is organised into three sequential and compulsory components:

I. Taught Component: Evidence-Based Public Health Policy & Practice (EBPHP) and Understanding Leadership, Management, & Organisation (ULMO)
II. Research Study I: Organisational & Policy Analysis (OPA) Project
III. Research Study II: Thesis project

Taught Components

The taught component consists of two compulsory modules. In the first module, Evidence-Based Public Health Policy (EBPHP), I learned how to assess, synthesize, and use research-based information to influence public health policy and practice in a range of settings. Most importantly, I learned policy theories and research methods useful for understanding and analysing different health sector organizations acting within their health policy environments. I also found that the application of theories of the health policy process (Agenda setting – Policy formulation – Policy implementation – Policy evaluation) was very helpful for better understanding and assessing how and why some health programs and policies could be particularly successful and effective.

In the second module, Understanding Leadership, Management, and Organizations (ULMO), I learned how to apply a range of organisational and business theories and tools to develop an understanding of the role of leaders, organisational management, and strategic planning for public health organizations. I was less experienced about the aspects of strategic planning and organisational changes in the field of public and global health. From this module, I particularly learned that the policy environment often changes, and it is important for public health organisations and leaders to make strategic decisions regarding organisational changes, in order to better position themselves to meet their objectives in a policy environment that is constantly changing.

Research Study I: Organisational & Policy Analysis (OPA) Project

The second component of the DrPH programme was the OPA Project, which allows students to observe and analyse the workings of a public health organisation in its policy environment and to gain a better understanding of how to develop effective public health organisations, influence public policy and deliver public health goals.

Before I joined the DrPH programme, I was involved in family planning and reproductive health projects in Ethiopia for several years. From my experiences, I was always interested in why birth rates were often high in many Sub-Saharan African (SSA) countries, leading to social and health problems such as limited access to education for women and maternal and child mortality, but family planning programs were not well implemented in SSA countries.

Therefore, my four-month fieldwork (July 2018 – November 2019) was conducted at the Population and Youth Section of the UN Economic Commission for Africa (UNECA), located in Addis Ababa, Ethiopia. My OPA project was entitled 'Generation of Political Priority of Family Planning in the United Nations Economic Commission for Africa (UNECA)', and this project was awarded a DrPH OPA Travelling scholarship. I was placed at the PYS as a research intern for four months, and I was involved in daily tasks at the PYS while conducting participant observations, document reviews, and key informant interviews for my OPA project.

I used both policy and organisational theories and tools, which I learned from the two compulsory modules from the DrPH programme. For the OPA report, I undertook a thematic analysis based on the McKinsey 7S model to examine how UNECA's organizational elements are aligned with each other and the Shiffman and Smith framework to assess the political priority of family planning on the agenda for the demographic dividend in UNECA. From my OPA project, I realised that health agendas can be politically driven and how UN bureaucracies and leadership could influence their member states and key stakeholders to generate political priority for family planning at the international, national, and organizational levels.

Research Study II: Thesis project

When I applied for the DrPH programme at LSHTM in 2017, my research proposal aimed to explore geographical variations in district-level fertility in Ethiopia by using Ethiopia's population and housing census data, since the census data is the only available data for districts. According to Ethiopia's constitution, the country is required to conduct a census every 10 years, and the last census was in 2007. Therefore, the fourth census was officially planned to be conducted in 2017.

During my OPA project in Ethiopia, I visited the Central Statistical Agency (CSA) of Ethiopia several times to obtain information about the fourth census's schedule. Unfortunately, I was unable to receive any updates. Afterwards, the Ethiopian government announced that the fourth census was postponed due to political issues in 2018, financial issues in 2019, and the COVID-19 pandemic in 2020. As a result, I realised that waiting for census data was not only a good idea for my thesis project, but also, the lack of local-data was a real-world and complex problem that could pose barriers to assessing and monitoring population and health in SSA countries.

Therefore, by the time I completed my OPA project, I strongly felt that I should learn Small Area Estimation using existing data to deal with the lack of local-data and to answer my research question in Ethiopia. The possible technique was to utilise spatial Bayesian models with R-INLA by using the Demographic and Health Survey (DHS) data. To hone my statistical and demographic skills, I took three relevant taught modules at LSHTM: Population Dynamics & Projection, Spatial Epidemiology in Public Health, and Survival Analysis and Bayesian Statistics. I also contacted several academic scholars who had used similar models in their published papers to get advice on applying R-INLA to my thesis project. Furthermore, my second and third supervisors changed from two reproductive health experts to one spatial statistician and one formal demographer.

Eventually, I could use spatial statistics models to explore the spatial aspects of geographical variations in district-level fertility in Ethiopia between 2000 and 2016. In contrast to the taught compulsory modules and the OPA project, the thesis project was a more independent work. However, thesis supervision at LSHTM could profoundly improve my academic writing and analytical skills and my ability to adapt scientific knowledge.

Overall, the DrPH programme gave me an opportunity to widen my views on both organisational aspects to assess the role of public health organisations and research aspects to present scientific evidence for evidence-informed decision making in health policy and practice. This experience will undoubtedly contribute to my future career in global health research and practice.

Appendix 1: Ethics Approval

Ethics approval

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
United Kingdom
Switchboard: +44 (0)20 7636 8636

www.lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Observational / Interventions Research Ethics Committee

Mr Myunggu Jung
LSHTM

19 March 2021

Dear Myunggu

Submission Title: Geographical fertility variations in Ethiopia

LSHTM Ethics Ref: 25580

Thank you for responding to the Observational Committee Chair's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Conditions of the favourable opinion

Approval is dependent on local ethical approval having been received, where relevant.

Approved documents

The final list of documents reviewed and approved is as follows:

Document Type	File Name	Date	Version
Investigator CV	CV_Myunggu Jung	03/03/2021	1
Local Approval	Email from ICF	11/03/2021	1
Local Approval	Local Approval_DHS2016	11/03/2021	1
Local Approval	Local Approval_DHS2011	11/03/2021	1
Local Approval	ICF_IRB_EDHS_2016	11/03/2021	1
Local Approval	ICF_IRB_EDHS_2011	11/03/2021	1
Local Approval	ICF_IRB_EDHS_2005	11/03/2021	1
Protocol / Proposal	Study Protocol_Ver_3	11/03/2021	3
Other	Research Ethics Training_Myunggu Jung	15/03/2021	2
Covering Letter	Cover Letter_25580	15/03/2021	1
Covering Letter	Cover Letter_25580_2	18/03/2021	2

After ethical review

The Chief Investigator (CI) or delegate is responsible for informing the ethics committee of any subsequent changes to the application. These must be submitted to the committee for review using an Amendment form. Amendments must not be initiated before receipt of written favourable opinion from the committee.

The CI or delegate is also required to notify the ethics committee of any protocol violations and/or Suspected Unexpected Serious Adverse Reactions (SUSARs) which occur during the project by submitting a Serious Adverse Event form.

An annual report should be submitted to the committee using an Annual Report form on the anniversary of the approval of the study during the lifetime of the study

At the end of the study, the CI or delegate must notify the committee using the End of Study form.

All aforementioned forms are available on the ethics online applications website and can only be submitted to the committee via the website at: http://leo.lshtm.ac.uk.

Further information is available at: www.lshtm.ac.uk/ethics.

Yours sincerely,

Page 1 of 2

[redacted]

Professor Jimmy Whitworth
Chair

ethics@lshtm.ac.uk
http://www.lshtm.ac.uk/ethics/

Improving health worldwide

Appendix 2: Sample sizes of Ethiopia demographic and health surveys (EDHS), 2000-2016

Number of primary sampling units (PSUs) and individuals from EDHS

a) 2000 EDHS

Region	2000 EDHS (A)		PSUs without GPS (B)		2000 Study (A - B)	
	PSU	Individual	PSU	Individual	PSU	Individual
Tigray	50	1,306	0	0	50	1,306
Afar	31	858	0	0	31	858
Amhara	78	1,909	0	0	78	1,909
Oromia	90	2,578	0	0	90	2,578
Somali	28	844	1	22	27	822
Benishangul- gumuz	38	992	0	0	38	992
SNNP	70	2,028	3	92	67	1,936
Gambela	35	876	1	24	34	852
Harari	33	908	0	0	33	908
Addis Ababa	51	2,015	0	0	51	2,015
Dire Dawa	35	1,053	1	36	34	1,017
Total	**539**	**15,367**	**6**	**174**	**533**	**15,193**

b) 2005 EDHS

Region	2005 EDHS (A)		PSUs without GPS (B)		2005 Study (A - B)	
	PSU	Individual	PSU	Individual	PSU	Individual
Tigray	50	1,257	1	21	49	1,236
Afar	34	789	0	0	34	789
Amhara	81	1,943	0	0	81	1,943
Oromia	87	2,230	0	0	87	2,230
Somali	30	669	3	56	27	613
Benishangul- gumuz	31	846	1	31	30	815
SNNP	84	2,087	2	44	82	2,043
Gambela	29	729	2	57	27	672
Harari	29	844	0	0	29	844
Addis Ababa	50	1,869	0	0	50	1,869
Dire Dawa	30	807	0	0	30	807
Total	**535**	**14,070**	**9**	**209**	**526**	**13,861**

c) 2011 EDHS

Region	2011 EDHS (A)		PSUs without GPS (B)		2011 Study (A - B)	
	PSU	Individual	PSU	Individual	PSU	Individual
Tigray	60	1,728	2	57	58	1,671
Afar	47	1,291	4	109	43	1,182
Amhara	73	2,087	2	63	71	2,024
Oromia	80	2,135	3	80	77	2,055
Somali	33	914	8	218	25	696
Benishangul- gumuz	49	1,259	2	75	47	1,184
SNNP	70	2,034	0	0	70	2,034
Gambela	46	1,130	5	162	41	968
Harari	42	1,101	0	0	42	1,101
Addis Ababa	54	1,741	0	0	54	1,741
Dire Dawa	42	1,095	1	21	41	1,074
Total	**596**	**16,515**	**27**	**785**	**569**	**15,730**

d) 2016 EDHS

Region	2016 EDHS (A)		PSUs without GPS (B)		2016 Study (A - B)	
	PSU	Individual	PSU	Individual	PSU	Individual
Tigray	63	1,682	0	0	63	1,682
Afar	53	1,128	0	0	53	1,128
Amhara	71	1,719	0	0	71	1,719
Oromia	75	1,892	3	74	72	1,818
Somali	66	1,391	16	314	50	1,077
Benishangul- gumuz	50	1,126	0	0	50	1,126
SNNP	71	1,849	0	0	71	1,849
Gambela	50	1,035	1	24	49	1,011
Harari	44	906	0	0	44	906
Addis Ababa	56	1,824	0	0	56	1,824
Dire Dawa	44	1,131	1	29	43	1,102
Total	**643**	**15,683**	**21**	**441**	**622**	**15,242**

Appendix 3: Fertility differences by residence and education

Fertility differentials by residential areas and educational levels from the most recent DHS reports

Country	Survey	Urban (A)	Rural (B)	No education or primary (C)	Secondary or higher (D)	urban-rural differentials (B-A)	educational differentials (C-D)
Ethiopia	**2016 DHS**	**2.3**	**5.2**	**5**	**2.1**	**2.9**	**2.9**
Cote d'Ivoire	2011-12 DHS	3.7	6.3	5.5	2.6	2.6	2.9
Zambia	2018 DHS	3.4	5.8	5.7	3.6	2.4	2.1
Mozambique	2018 MIS	3.9	6.2	6.1	3.5	2.3	2.6
Liberia	2013 DHS	3.8	6.1	5.5	3.4	2.3	2.1
Senegal	2018 DHS	3.2	5.5	5.1	2.8	2.3	2.3
Tanzania	2015-16 DHS	3.8	6	5.7	3.6	2.2	2.1
Cameroon	2018 DHS	3.8	6	5.9	3.8	2.2	2.1
Liberia	2019-20 DHS	3.4	5.5	5.1	3.2	2.1	1.9
Gambia	2019-20 DHS	3.9	5.9	5.5	3.4	2	2.1
Togo	2013-14 DHS	3.7	5.7	5.5	3.5	2	2
Congo	2011-12 DHS	4.5	6.5	6.6	4.5	2	2.1
Sierra Leone	2019 DHS	3.1	5.1	5	3	2	2
Uganda	2016 DHS	4	5.9	6	4.2	1.9	1.8
Congo Democratic Republic	2013-14 DHS	5.4	7.3	7.5	5.6	1.9	1.9
Mali	2018 DHS	4.9	6.8	6.7	4.5	1.9	2.2
Burkina Faso	2017-18 MIS	3.7	5.6	5.7	3.3	1.9	2.4
Senegal	2019 DHS	3.8	5.6	5.4	3.6	1.8	1.8
Haiti	2016-17 DHS	2.1	3.9	4.3	2.2	1.8	2.1
Zimbabwe	2015 DHS	3	4.7	5.1	3.7	1.7	1.4
Guinea	2018 DHS	3.8	5.5	5.1	3.5	1.7	1.6
Malawi	2015-16 DHS	3	4.7	4.9	3.2	1.7	1.7
Ghana	2016 MIS	3.4	5.1	5.4	3.6	1.7	1.8
Burundi	2016-17 DHS	4.1	5.7	6.1	4.1	1.6	2
Lesotho	2014 DHS	2.3	3.9	3.9	2.9	1.6	1
Madagascar	2016 MIS	2.7	4.3	4.6	3.3	1.6	1.3
Ghana	2019 MIS	3.1	4.6	4.9	3.3	1.5	1.6
Nigeria	2018 DHS	4.5	5.9	6.5	4.2	1.4	2.3

Kenya	2014 DHS	3.1	4.5	4.7	3	1.4	1.7
Chad	2014-15 DHS	5.4	6.8	6.7	4.8	1.4	1.9
Comoros	2012 DHS	3.5	4.8	5.5	3.1	1.3	2.4
Guatemala	2014-15 DHS	2.5	3.7	3.8	2.2	1.2	1.6
Timor-Leste	2016 DHS	3.5	4.6	4.7	4	1.1	0.7
Tajikistan	2017 DHS	3	4	4	3.8	1	0.2
Pakistan	2017-18 DHS	2.9	3.9	4.1	2.9	1	1.2
Honduras	2011-12 DHS	2.5	3.5	3.6	2.4	1	1.2
Kyrgyz Republic	2012 DHS	3	4		3.6	1	-
Egypt	2014 DHS	2.9	3.8	3.7	3.5	0.9	0.2
Nepal	2016 DHS	2	2.9	3.1	1.9	0.9	1.2
Benin	2017-18 DHS	5.2	6.1	6.1	4.2	0.9	1.9
Colombia	2015 DHS	1.8	2.6	3.1	1.9	0.8	1.2
Papua New Guinea	2016-18 DHS	3.5	4.3	4.4	3.7	0.8	0.7
Cambodia	2014 DHS	2.1	2.9	3.1	2.3	0.8	0.8
South Africa	2016 DHS	2.4	3.1	3.4	2.6	0.7	0.8
Maldives	2016-17 DHS	1.8	2.5	2.4	2.2	0.7	0.2
Rwanda	2014-15 DHS	3.6	4.3	4.5	3	0.7	1.5
India	2015-16 DHS	1.8	2.4	2.9	1.9	0.6	1
Jordan	2012 DHS	3.4	3.9	3.6	3.5	0.5	0.1
Myanmar	2015-16 DHS	1.9	2.4	2.8	1.9	0.5	0.9
Philippines	2017 DHS	2.4	2.9	4.1	2.5	0.5	1.6
Indonesia	2012 DHS	2.4	2.8	2.9	2.6	0.4	0.3
Jordan	2017-18 DHS	2.7	3.1	3.3	2.7	0.4	0.6
Bangladesh	2017-18 DHS	2	2.3	2.5	2.2	0.3	0.3
Dominican Republic	2013 DHS	2.4	2.6	3.3	2.3	0.2	1
Albania	2017-18 DHS	1.7	1.9	2.4	1.7	0.2	0.7
Armenia	2015-16 DHS	1.7	1.8	2.7	1.7	0.1	1

Appendix 4: Median age at first marriage and sexual intercourse

Differentials between (a) median age at first marriage and (b) at first sexual intercourse by eleven regional states in Ethiopia, 2000-2016

	2000		2005		2011		2016	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
National	16	16	16.1	16.1	16.5	16.6	17.1	16.6
Tigray	15.6	15.5	15.6	15.5	16.6	15.7	16.6	16.1
Afar	15.5	15.5	16.4	16.1	16.5	16.9	16.4	16.2
Amhara	14.3	14.4	14.2	14.6	14.7	15.1	15.7	15.5
Oromia	16.4	16.5	16.7	16.9	16.9	17	17.2	16.7
Somali	17.3	17.4	18	18.4	17.6	17.9	18.1	17.9
Benishangul Gumuz	15.6	15.6	15.3	15.6	15.7	16	16.8	16.5
SNNPR	17.7	17.7	17.2	17.3	17.9	17.9	17.7	17.8
Gambela	16.2	16	15.7	15.7	17.1	16.9	16.9	16.2
Harari	16.8	16.4	18.6	18.6	17.7	17.9	18.3	17.7
Addis Ababa	19.3	18.2	21.9	20	21.4	19.5	23.9	20.4
Dire Dawa	19.4	18.7	17.8	17.5	19	19.3	18.1	17.7

Appendix 5: Shapiro-Wilk Test for total fertility rate, the median age at first marriage and the ethnolinguistic diversity index

Shapiro-Wilk Test for total fertility rate, the median age at first marriage and the ethnolinguistic diversity index.

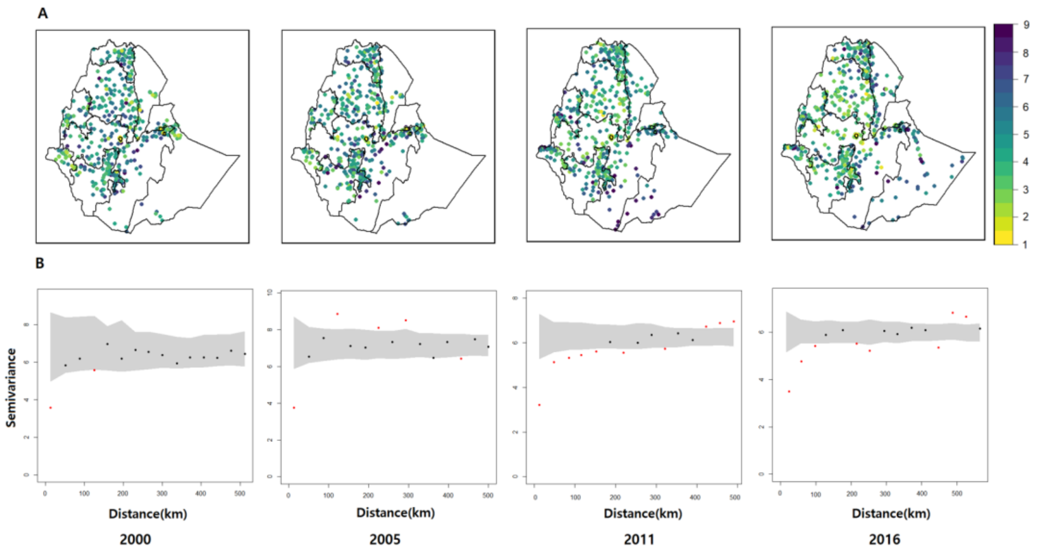
	2000		2005		2011		2016	
	W	**P-value**	**W**	**P-value**	**W**	**P-value**	**W**	**P-value**
Total fertility rate	0.986	0.432	0.974	0.214	0.964	0.132	0.972	0.213
Median age at first marriage	0.996	0.227	0.995	0.132	0.994	0.094	0.993	0.087
Ethnolinguistic diversity index	0.955	0.561	0.936	0.345	0.948	0.387	0.952	0.547

Appendix 6: Primary sampling unit level observations and variograms of study variables

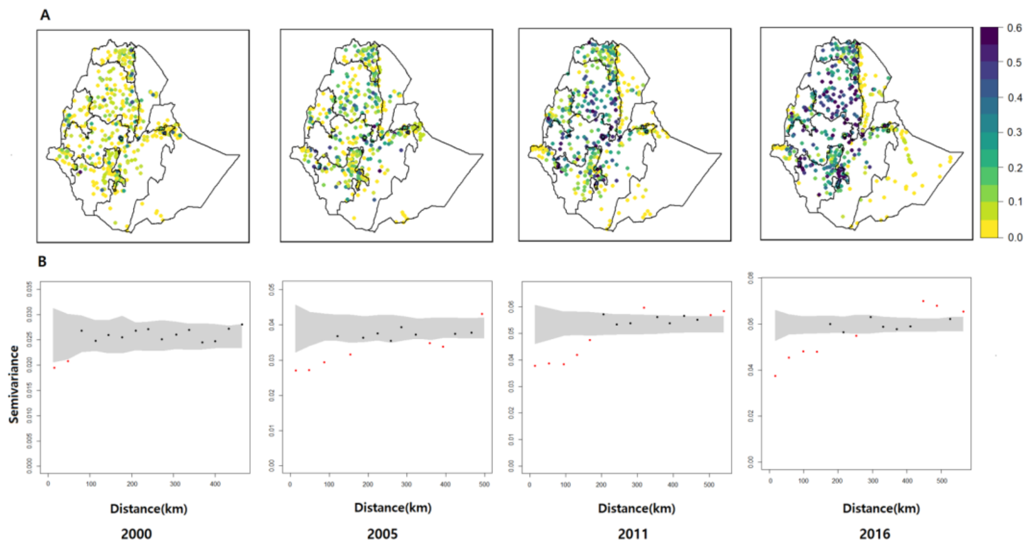
Primary sampling unit (PSU)-level observations (A) and variograms (B) of six study variables from Ethiopia Demographic and health survey (EDHS), 2000-2016: 1) total fertility rate, 2) modern contraceptive prevalence, 3) median age at first marriage, 4) proportion of women living in urban area, 5) proportion of women having secondary education, 6) index of ethnolinguistic diversity. I compared the empirical semi-variograms with a Monte Carlo envelope of empirical semi-variograms computed from random permutations of the data holding the locations fixed. If the empirical semi-variogram lies outside the Monte Carlo envelope, there is evidence of spatial correlation.

The empirical variograms (B) are represented by dots, where black dots indicates that variogram estimates are within the 95% envelope, while red dots reflect variogram estimates outside the 95% envelope. The Monte Carlo envelope (gray shading) displays pointwise 95% coverage of 1,000 permutations. Semi-variograms show that there are evidence of spatial correlation since the observed variograms of all study variables lie partly outside the 95% pointwise envelope between 2000 and 2016.

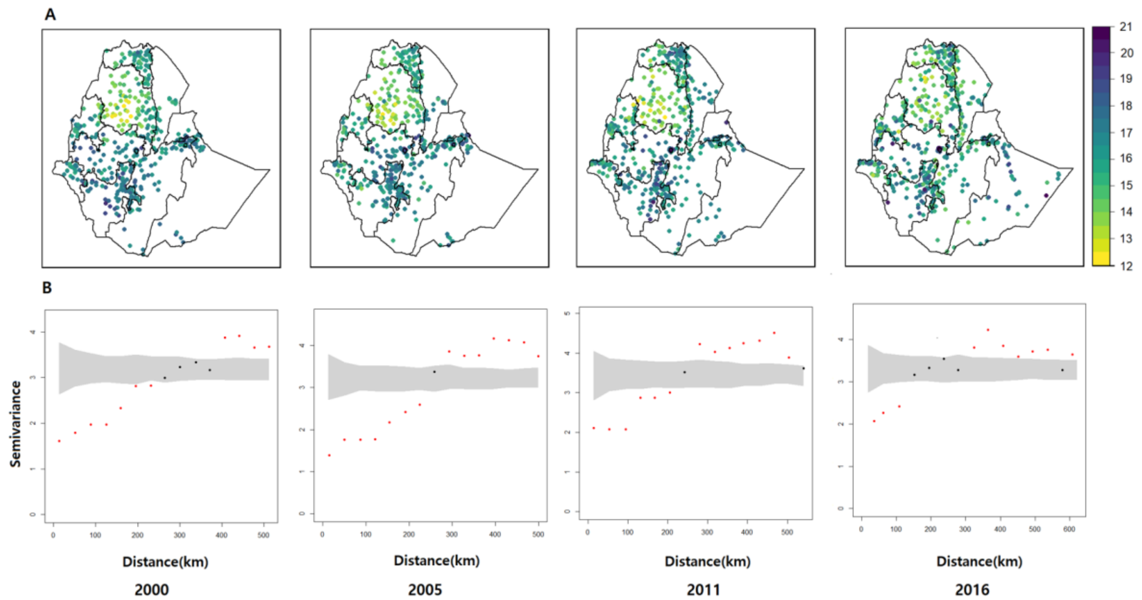
1) Total fertility rate



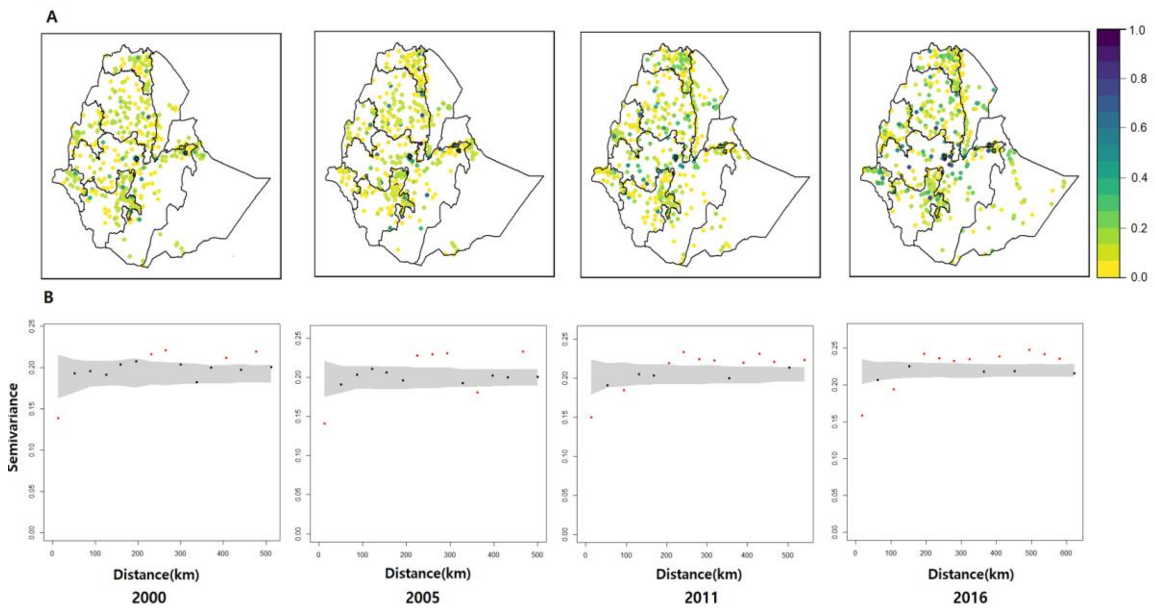
2) Modern Contraceptive Prevalence (mCP) among married women of reproductive age



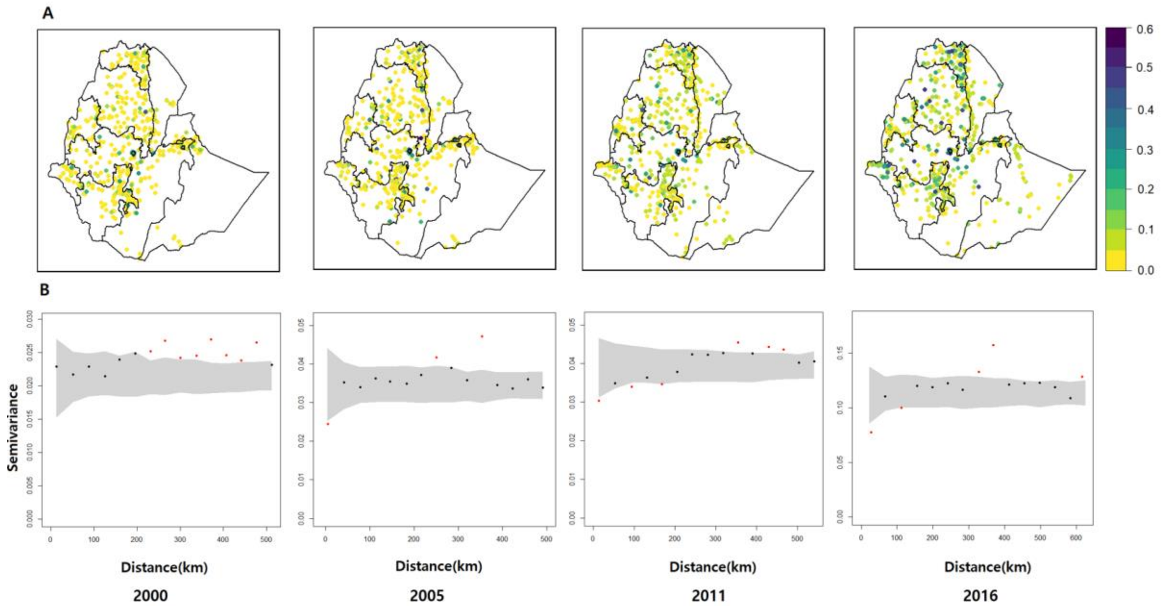
3) Median age at first marriage



4) Proportion of women living in urban areas



5) Proportion of women having secondary education



6) Index of ethnolinguistic diversity