Analysis of Childhood Diseases and Malnutrition in Developing Countries of Africa

Khaled Khatab

München, 6. March 2007

Analysis of Childhood Diseases and Malnutrition in Developing Countries of Africa

Khaled Khatab

Dissertation
zur Erlangung des Grades Doctor oeconomiae publicae
(Dr. oec. publ.)
an der Ludwig–Maximilians–Universität, München

vorgelegt von
Khaled Khatab
2007

Institut für Statistik, Fakultät für Mathematik, Informatik und Statistik

Referent: Prof.Dr Ludwig Fahrmeir

Korreferent: Prof. Dr Joachim Winter

Promotionsabschlussberatung: 2007
Rigorosum 9. July 2007

Contents

Acknowledgment

I would like to express my gratitude to my supervisor, Prof. Dr. Ludwig Fahrmeir, for his unflinching and generous supervision as well as easy accessibility at all times. His constructive criticism and suggestions have helped me in widening my research abilities. In spite of his demanding schedule, he made himself available to me at any time for useful discussions, which helped in finalizing this work. Furthermore, I was also so happy to work with him during my research period and during my residence period in Germany, either under his chair directly or under the "*Deutsch Forschungsgemeinschaft Sonderforschungsbereich*" (SFB) 386. I am also grateful to Prof. Dr. Winter, who agreed to be a co-supervisor for this work, and made himself available to me any time I needed his suggestions and his help.

My thanks are due to my colleagues and my friends Dr. David Rummel, Dr. Andreas Brezger, Ralf Breuninger, Sven Steinert, who have contributed their time, useful suggestions and efforts. I would like to thank Dr. Alexander Raach, who helped me in working with the MCMC package.

Also, I have to acknowledge the support I received from SFB 386 during my research period.

I am also grateful to my friend Dr. Kandala, of the University of Warwick, for his insightful comments, helpful advice and his discussions which helped me very much in editing this work. To him I also say a special *thank you.* My thanks go to Dr. Samson for his significant discussions. Furthermore, I must thank Prof. Stefan Lang for his collaboration and his useful suggestions.

I would like to express my deep gratitude to my friends Khaled Mahmmod and Mohammed Talal for their dear friendship and their supporting to me during my residence period in Germany. To them I really must say *many thanks*. Also, I am again grateful to my friend Mohamed El salhi for his useful suggestions, support and his help with many technical problems related to this work.

My thanks go to my friend Farooq Bashir for his supporting to me during the last few months of my research period.

I cannot begin to acknowledge my princess, Amie. Thanks for her help, her understanding and her painstaking support during this stressful time. My very special gratitude goes to her.

Finally, I really would like to express my deep gratitude and appreciation to my **Mother** and my **Father** who suffer so much for me. I am deeply indebted to them for their patience, concern and love. I cannot praise them enough for their support. I would also like to thank my brother and my sister for their support. To them all I have dedicated this dissertation.

GOD be praised for guiding us. We could not possibly be guided, if it were not that GOD has guided us.

Introduction

Child disease and malnutrition reflect a country's level of socio-economic developing and quality of life.

This thesis is an empirical work dealing with childhood disease and malnutrition in African developing countries, particularly, in Egypt and Nigeria.

The objective of this work is to examine the impact of socioeconomic and public health factors on childhood diseases and malnutrition in mentioned countries. The causes of child's illness or child's undernutrition are multiple. This work focuses on some risk factors which are assumed to cause the child's diseases and malnutrition as suggested by some previous works (see Kandala, 2001; Adebayo, 2002). Our analysis started with a large number of covariates including a set of bio-demographic and socioeconomic variables, such as current working status of mothers, place of residence, access to toilet facilities, etc (see chapter 2). The analyses are based on data from the 2003 household survey for Egypt and Nigeria for the Demographic and Health Surveys (DHS). More details about the data set are mentioned in the first chapter. The statistical analysis in this thesis is based on modern Bayesian approaches which allow a flexible framework for realistically complex models. These models allow us to analyze usual linear effects of categorical covariates, nonlinear effects of continuous covariates and the geographical effects within a unified semi-parametric Bayesian framework for modelling and inference. A first step of this work is to analyze the effects of the different types of covariates on response variables, diarrhea, fever, and cough which represent the child's diseases in our application. In this step, a Bayesian geoadditive logit model for binary response variables is used (see

Fahrmeir and Lang, 2001). In a second step, we employ separate geoadditive probit models (instead of logit models used in the previous step) to the binary listed variables. Based on the results of the separate analyses, we applied geoadditive latent variable probit models (recently suggested by Raach, 2005; Raach and Fahrmeir, 2006) where the three observable disease variables are assumed to be indicators for the latent variable "health status" for the children. In this step, we also compared the results of the separate geoadditive probit models with the results of the latent variable models. As a third step, we used geoadditive Gaussian regression and latent variable models to analyze the malnutrition status of children in both countries. Finally, we used latent variable models for diseases and nutrition indicators together. In the final step, models with one as well as with two latent variables have been estimated using mixed indicators (binary indicators "health status", and continuous indicators "nutrition status") and the results are compared.

The analyses in this work are based on semi-parametric models developed by Fahrmeir and Lang (2001) and Brezger and Lang (2005), and on geoadditive latent variable models, recently suggested by Raach (2005) and Fahrmeir and Raach (2006). All computations to implement the methodology discussed here are carried out with BayesX program-version 1.4 (Brezger, Kneib, and Lang, 2005), and with R using the MCMC package (Raach, 2005 and Fahrmeir and Raach, 2006). All empirical results are discussed at the end of the relevant chapters.

Chapter 1

Descriptive and Explanatory Analysis of Variables

1.1 Data Set and Methods

The analyses in this thesis are based on data available from the 2003 DHS. The DHS uses standard survey instruments to collect data on household members such as sex of child, age of child, mother's age, current employment status of mother, mother's educational attainment, exposure to mass media, the type of toilet facility etc. It collects information on household living conditions such as housing characteristics, on childhood morbidity, malnutrition and child health from mothers in reproductive ages (15-49). The data is based on national samples that have been collected using questionnaires and allows for breakdowns by urban-rural and major regions and governorates.

With regard to measures of child health for children under 5 years, the focus in this work and in the analysis will be on the following: (1) Child morbidity such as a prevalence of diarrhea, fever and cough with difficulty of breathing (a symptom of respiratory infection) and (2) Child nutritional status with prevalence of malnutrition.

1.2 Childhood Disease

The diseases of children included in this work for Egypt and Nigeria are diarrhea, cough, and fever. These diseases are still a major cause of mortality among children in many developing countries, particularly in Sub-Saharan Africa. Yet, except for some descriptive reports by National Statistics Offices of these countries, few systematic studies of factors that influence the prevalence of diarrhea, cough and fever among young children were carried out in these countries. The success of health care intervention depends on a correct understanding of the socioeconomic, environmental and cultural factors that determine the occurrence of diseases, undernutrition and deaths. The mapping of variation in risk of child morbidity and child malnutrition can help in improving the targeting of scarce resources for public health interventions. Bearing in mind that direct mapping of relevant environmental risk factors (which may vary considerably in both space and time) is difficult and this has led to investigations of environmental proxies (Kandala et al., 2001). Our focus in next subsection is on the diseases which are used as response variables in this work.

Diarrhea

There is a variety of microorganisms that could be the main cause of the diarrhea disease, microorganism including viruses, bacteria and protozoans. Diarrhea affects the health of persons and causes loss of water and electrolytes as well a leading cause of both dehydration and death in some other cases. It is a public health problem related to water and sanitation. About 4 billion cases of diarrhea cause 2.2 million deaths, annually mostly among children under five (UNICEF, 2002). In the 2003 DHS, mothers were asked whether any of their children under five years of age had diarrhea at any time during the two-week period prior to the survey. We assumed diarrhea to be a binary variable, which is 1 when the child had disease and 0 if not. The same has done for fever and cough.

An example for percentage of children under five years of age who had diarrhea, cough or fever in two week preceding the survey with selected background characteristics for Egypt and Nigeria shown in tables 1.1, 1.3, and 1.4, respectively. They indicate the percentages of children under five years of age who had diseases at some time (May-June 2003) during two-week period before the survey.

Tables indicate that the children are in a higher risk of diseases during the first 20-24 months of age in both countries. The children living in rural areas are more likely to have diseases compared to their counterparts in urban areas.

Cough

Cough and difficult breathing are common problems among young children. The recent literature indicates the breastfed child who has a cough or cold may have difficulty feeding, however breastfeeding could be helped to fight the diseases. Along with diarrhea, acute respiratory infection (ARI), particularly pneumonia, is a common cause of death among infants and young children (DHS 2003). The prevalence of ARI has been estimated in the 2003 DHS by asking mothers if their children under five years of age had an illness with coughing accompanied by short rapid breathing in the two weeks before the survey. Disease of cough and short rapid breathing are symptoms of pneumonia, and thus the results of DHS are less appropriate for use in assessing the presence of other ARI-related conditions (cough and colds, wheezing, ear infection and streptococcal sore throat).

Fever

Most fevers in babies and children are caused by a viral (germ) infection. However, fever is less common and high fevers are unusual in young infants, and any fever should be considered a danger sign of very severe disease. The causes of fever could be as the next:

- An infection caused by germs called virus, parasites, or bacteria.
- Vaccinations, or immunization shots.

Table 1.1: Percentage of children under five years of age who had diarrhea in two week preceding the survey, selected background characteristics, Egypt and Nigeria, 2003).

characteristic	Percentage of child (had diarrhea)	Number of cases with diarrhea
Child's age-Egypt		
under 6 months	12	152
$6-11$months	20	270
$12-23$ months	30	401
$24-35$ months	19.8	265
$36-47$ months	12	160
$48-59$ months	6.2	87
Child's age-Nigeria		
under 6 months	8.5	79
$6-11$months	17.7	165
$12-23$ months	27	250
$24-35$ months	23.7	220
$36-47$ months	15.2	141
$48-59$ months	7.85	74
Sex-Egypt		
Male	56.2	751
female	43.8	584
Sex-Nigeria		
Male	52.1	484
female	47.9	445
Residence-Egypt		
Urban	30	397
Rural	70	938
Residence-Nigeria		
Urban	23.3	217
Rural	76.7	712
Place of residence-Egypt		
Urban Governorates	14	186
Lower Egypt	43	574
Urban	23.6	136
Rural	76.4	438
Upper Egypt	43	575
Urban	22	127
Rural	78	448
Region-Nigeria		
North Central	11.5	107
Northeast	42.9	398
Northwest	34	316
Southeast	2.9	27
South	5.7	52
Southwest	3	29

Variable	Obs	Mean	Std. Dev.	0:had no diseases	1:had diseases
Diarrhea-Nigeria	5186	0.179	0.383	4.257(82.09)	929(17.91)
Fever-Nigeria	5186	0.309	0.462	3.583(69.09)	1.603(30.91)
Cough-Nigeria	5186	0.235	0.424	3.967(76.49)	1.219(23.51)
Diarrhea-Egypt	6348	0.210	0.407	5.013(78.97)	1.335(21.03)
Fever-Egypt	6348	0.323	0.467	4.297(67.69)	2.051(32.31)
Cough-Egypt	6348	0.255	0.4361	4.725(74.43)	1.623(25.57)

Table 1.2: Overview of diseases in Egypt and Nigeria

- Sometimes children have a fever for no apparent reason.

1.2.1 Childhood Malnutrition

Childhood undernutrition is amongst the most serious health issues facing developing countries. It is an intrinsic indicator of well-being, but it is also associated with morbidity, mortality, impaired childhood development, and reduced labor productivity (Sen, 1999; UNICEF, 1998; Pritchett and Summers, 1994; Pelletier; 1998, Svedberg 1999).

To assess nutritional status, the 2003 DHS obtained measurements of height and weight for all children with the most of research focused on children below six years of age. Researchers distinguish between three types of malnutrition; wasting or insufficient weight for height indicating acute malnutrition; stunting or insufficient height for age indicating chronic malnutrition; and underweight or insufficient weight for age which could be a result of both stunting and wasting. Wasting, stunting, and underweight for a child i are typically determined using a Z-score which is defined as:

$$Z_i = \frac{AI_i - MAI}{\sigma}, \tag{1.1}$$

where AI refers to the individual anthropometric indicator (e.g. height at a certain age), MAI refers to the median of a reference population, and σ refers to the standard deviation of the reference population. The reference standard typically used for the calculation is the NCHS-CDC Growth Stan-

Table 1.3: Percentage of children under five years of age who had cough in two week preceding the survey, selected background characteristics, Egypt and Nigeria, 2003).

characteristic	Percentage of child (had cough)	Number of cases with cough
Child's age-Egypt		
under 6 months	8	131
$6-11$months	14.7	239
$12-23$ months	24	388
$24-35$ months	19.8	309
$36-47$ months	18.5	301
$48-59$ months	15.8	255
Child's age-Nigeria		
under 6 months	11	134
$6-11$months	17	206
$12-23$ months	23.4	285
$24-35$ months	19.8	242
$36-47$ months	17	206
$48-59$ months	11.3	146
Sex-Egypt		
Male	55.4	898
female	44.6	725
Sex-Nigeria		
Male	52.5	614
female	47.5	605
Residence-Egypt		
Urban	37	600
Rural	63	1023
Residence-Nigeria		
Urban	34.3	418
Rural	65.7	801
Place of residence-Egypt		
Urban Governorates	14	226
Lower Egypt	36	583
Urban	11	177
Rural	25	406
Upper Egypt	50	814
Urban	12	197
Rural	38	617
Region-Nigeria		
North Central	16.5	201
Northeast	32.3	394
Northwest	21	256
Southeast	9.4	115
South	12	146
Southwest	8.8	107

Table 1.4: Percentage of children under five years of age who had fever in two week preceding the survey, selected background characteristics, Egypt and Nigeria, 2003).

characteristic	Percentage of child (had fever)	Number of cases with fever
Child's age-Egypt		
under 6 months	9.3	191
$6-11$months	16	329
$12-23$ months	26	534
$24-35$ months	20.8	415
$36-47$ months	15.75	323
$48-59$ months	12.63	259
Child's age-Nigeria		
under 6 months	9	149
$6-11$months	16.2	261
$12-23$ months	25.4	408
$24-35$ months	20.4	329
$36-47$ months	16.4	263
$48-59$ months	11	177
Sex-Egypt		
Male	54.5	1119
female	46.5	932
Sex-Nigeria		
Male	50.8	818
female	49.2	785
Residence-Egypt		
Urban	34.7	712
Rural	56.3	1339
Residence-Nigeria		
Urban	32	516
Rural	68	1090
Place of residence-Egypt		
Urban Governorates	13.4	275
Lower Egypt	31.6	648
Urban	8.6	177
Rural	23	471
Upper Egypt	55	1128
Urban	12.7	260
Rural	42.3	868
Region-Nigeria		
North Central	13.4	215
Northeast	28.3	455
Northwest	33.2	535
Southeast	8.6	138
South	9.3	150
Southwest	7.2	110

dard that has been recommended for international use by WHO. Each of the indictors measures somewhat different aspects of nutritional status.

Stunting

Stunting is an indicator of linear growth retardation relatively uncommon in the first few months of life. However it becomes more common as children get older. Children with *height-for-age* z-scores below minus two standard deviations from the median of the reference population are considered short for their age or stunted. Furthermore, children with z-scores below minus three standard deviations from the median of the reference population are considered to be severely stunted, while children with z-scores between minus three and minus two standard deviations are known to be moderately. In our application, however, we will only distinguish between children who are undernourished and those who are not.

Wasting

Wasting indicates body mass in relation to body length. Children whose *weight-for-height*'s z-scores are below minus two standard deviations (z-scores $< -2SD$) from the median of the reference population are considered wasted (i.e. too thin for their height) which implies that they are acutely undernourished otherwise they are not wasted. Whilst those with a z-score below -3 are considered severely wasted. Wasting results from either a lack of the ability to receive adequate nutrition shortly before the survey or an evidence of recent illness such as diarrhea which causes loss of weight and consequently results in a start of malnutrition.

Underweight

Underweight is a composite index of stunting and wasting. This means children may be underweight if they are either stunted or wasted, or both. In a similar manner to the two previous anthropometric incidences, children may be underweight when their z-score is below minus two standard deviations and they are severely or moderately so if their z-score is lower than two standard deviations. Our application focuses on the three indicators of malnutrition status, but we use the z-score (in a standardized form) as continuous variable.

1.3 Descriptive and Explanatory Analysis of Variables

We will go through the description and explanation of the variables used in this thesis. This description has to be for the countries Egypt and Nigeria; those would be included in our application. The variables that will be used in this analysis will be described in this section to asses the most important influential factors on child diseases and malnutrition. In this study the following covariates were included.

1.3.1 Spatial Covariates

The information of the geographical location (region or governorate) where the illness or the undernourished child lives at the time of interview is a significant contribution of the DHS data set to an understanding of the child disease and undernutrition status in both Egypt and Nigeria. The information has been used (but not widely) by some previous studies on African child nutritional status (see Kandala, 2001; Adebayo, 2002) but is rarely used in the case of child disease.

In the case of Egypt, there are 20 governorates included. For Nigeria, 37 regions apply. Figure 1.4 (right) shows that Lower Egypt and essentially some districts in Nile Delta are associated with significantly high rate of illness and the left panel suggests that the diarrhea disease is significantly higher in some districts of the central region and in some districts in southern of Nigeria. The red area indicates a negligible effect for this disease within these areas, while the green area reflects a strong effect in these regions and the yellow area indicates that there are almost no cases found in these regions.

Figures 1.5 and 1.6 (right panel) suggest that cough and fever disease are significantly higher in districts around the Nile Delta in Egypt and on the other hand the left panel of Figure 1.5 shows that the cough is significantly higher in southwestern and some northwestern districts.

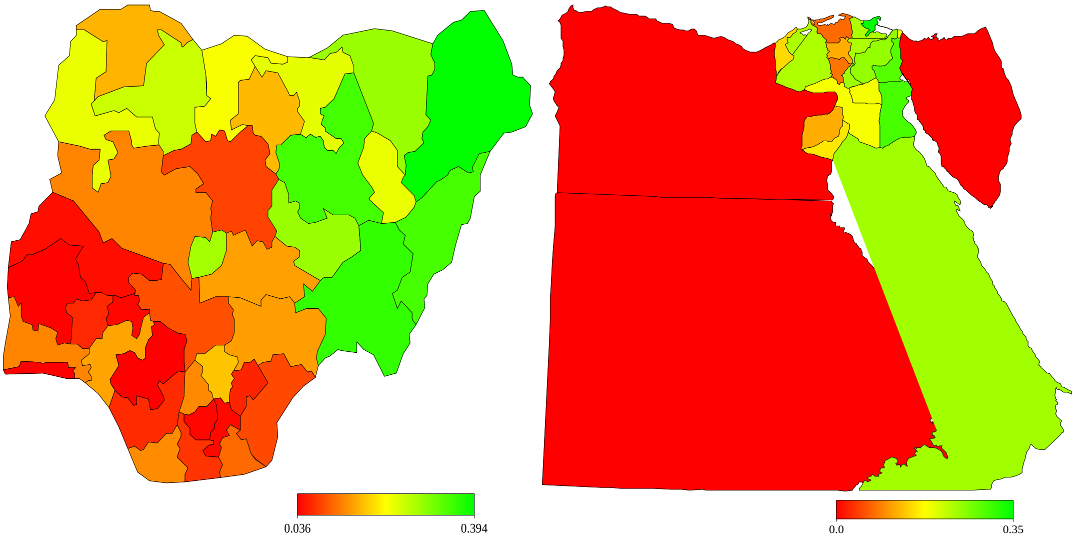


Figure 1.1: Rate of diarrhea in Nigeria (left) and in Egypt (right).

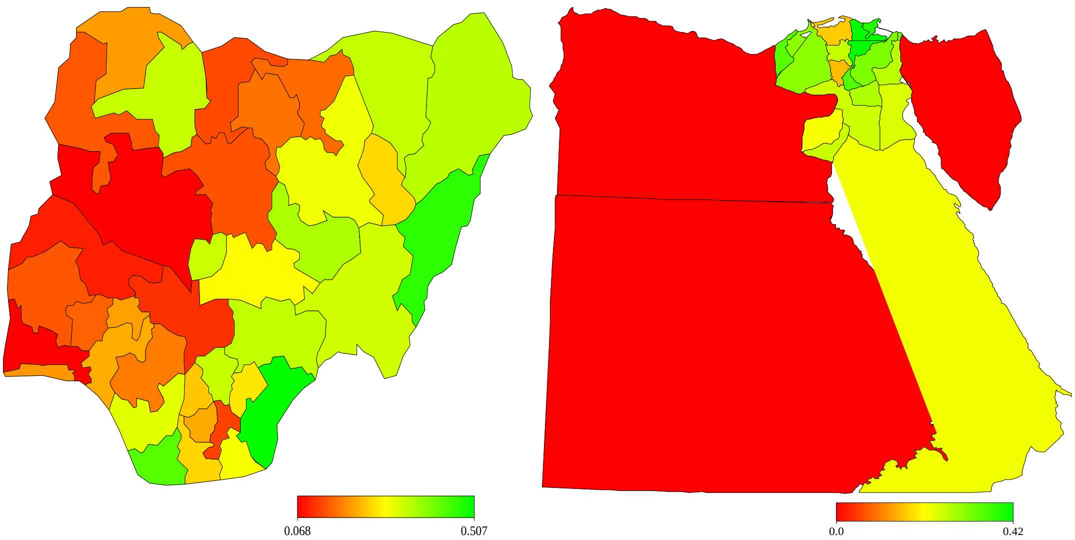


Figure 1.2: Rate of cough in Nigeria (left) and in Egypt (right).

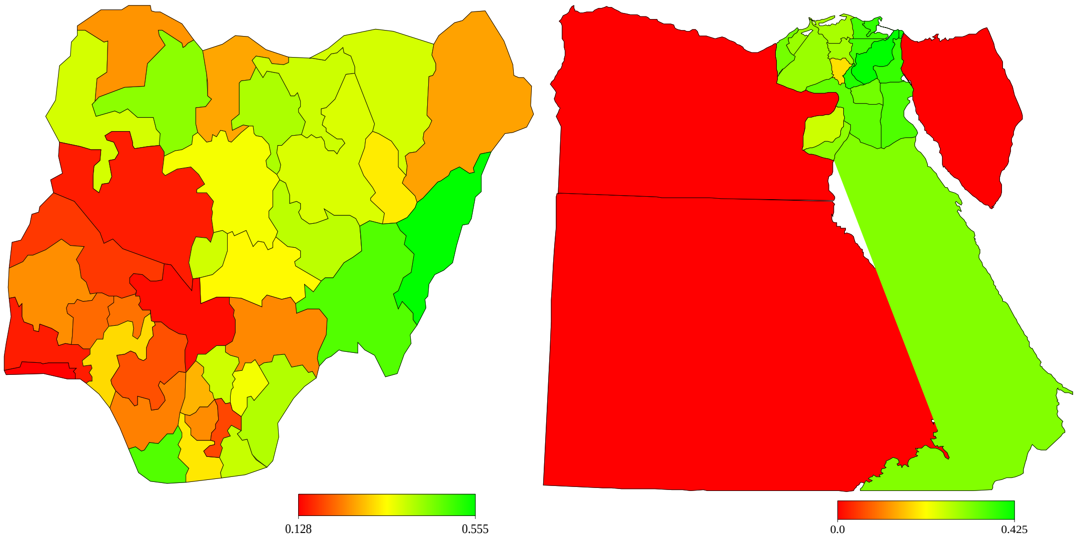


Figure 1.3: Rate of fever in Nigeria (left) and in Egypt (right).

1.3.2 Metrical Covariates

Child's age

The prevalence of diseases and stunting rises with age. According to the World Health Organization (WHO), children should receive the complete schedule of recommended vaccinations by 12 months of age. In Nigeria, only 13 percent of children age 12-23 months are fully immunized. However, in Egypt, virtually all children 12-23 months have received at least some of the recommended vaccinations and an overall 88 percent of children are considered as immunized against all major preventable childhood diseases (DHS 2003).

BMI Body Mass Index

BMI is a tool for indicating weight status in adults. The risk of some diseases increases as BMI increases.

The effect of the mother's body mass index, defined as the weight in kilograms divided by the square of height in meters. This effect can be explored by a non-parametrical function.

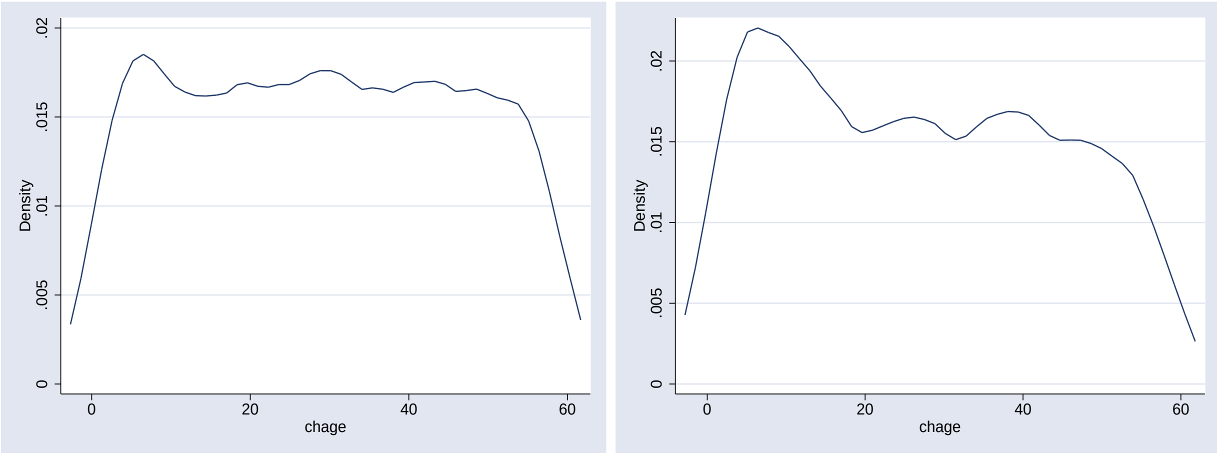
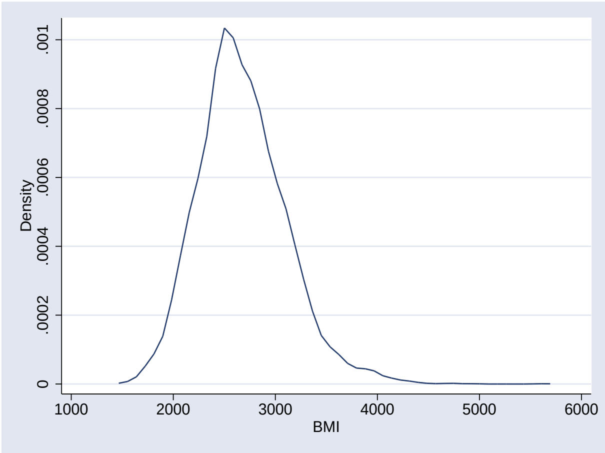


Figure 1.4: Kernel density estimates of child's age in Egypt (left) and Nigeria (right).

Mother's age at child's birth

The effect of the mother's age at child's birth may be explored by categorizing the three age groups respectively as in some previous studies: young mothers (less than 22 years old), middle-age group (between 22-35 years old), and old age group (greater than 35 years old). However, in our application we include this covariate as a metrical covariate to have more reasonable results. In Nigeria, more than 68 percent of all women are currently married, 25 percent have never been married, while negligible proportions of women are divorced or separated (3 percent) or widowed (2 percent). In Egypt, 92 percent of those interviewed are currently married, while 5 percent are widowed and 3 percent were either divorced or separated.



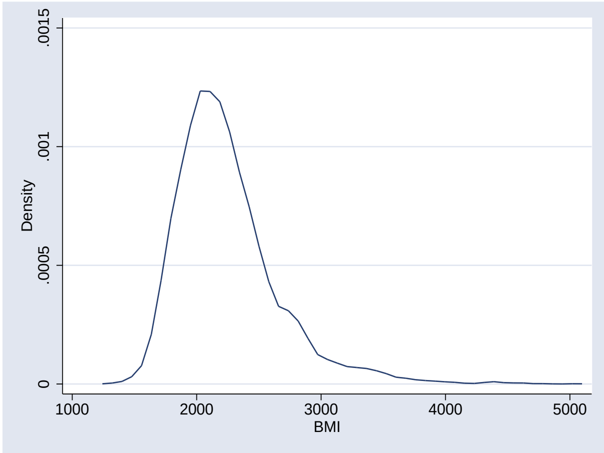


Figure 1.5: Kernel density estimates of mother's body mass index in Egypt (left) and Nigeria (right).

1.3.3 Categorical Covariates

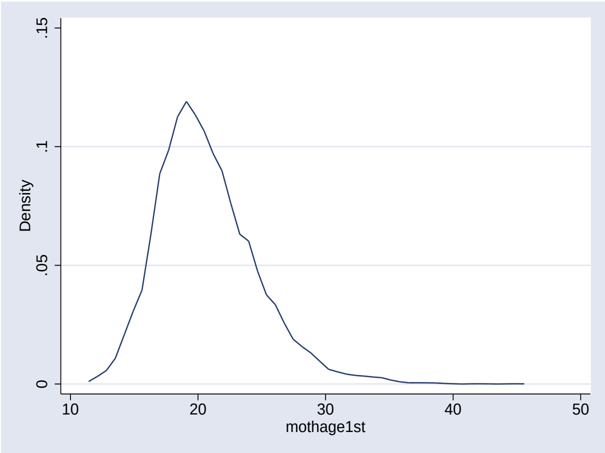
Current employment status of mother

Respondents who are currently employed or worked within the year before the survey were asked to state their occupation.

In our application, we distinguished between respondents who are currently working and respondents who are not working (reference category). The report by the surveys focuses on whether the mother was working at the time of the survey. Only 15.9 percent of those in the 2003 EDHS work for cash and overall 84 percent of women are not working or are not paid for work they do.

Mother's educational attainment

In this thesis mother's educational attainment is recorded into three categories: "no education and incomplete primary education "(reference category), "complete primary education and incomplete secondary education" and "complete secondary education and higher", respectively and in the latter analysis these categories are reduced to two categories: "incomplete secondary education" and "complete secondary school or higher" (reference



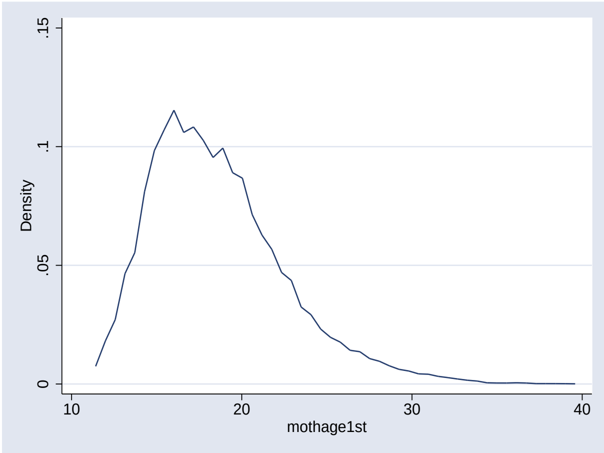


Figure 1.6: Kernel density estimates of mother's age in Egypt (left) and Nigeria (right).

category).

Birth interval

The period of time between two successive live births is referred to as a birth interval. Some research has shown that children born soon after a previous birth are at greater risk of illness and death than those born after a long interval. A short birth interval is defined to be no longer than 24 months (reference category) and it is associated with high morbidity, exhausting the mother.

Rural and urban residence

Place of residence, whether urban or rural, is determined by the interviewer according to the location of the interview. With regard to residence, 60 percent of women in Nigeria are from the rural areas while 40 percent are from urban areas, while 75 percent of the 2003 EDHS respondents live in rural areas and 43 percent are urban residents. In this application rural area is assumed to be the reference category.

Child's sex

Gender discrimination is found to have an effect on the disease of children

Factor	Egypt (n (%))	Nigeria (n (%))	coding
Place of residence			
Urban	2237(33.58%)	2118(35.13%)	1
Rural	4424(66.42%)	3911(64.87%)	-1.ref
Child's sex			
Male	3487(52.35%)	3062(50.79%)	1
Female	3174(47.65%)	2967(49.21%)	-1.ref
Working			
Yes	1209(18.15%)	3835(63.61%)	1
No	5452(81.85%)	2172(36.39%)	-1.ref
Mother's education			
No,			
Incomp.prim,			
Comp.prim,			
Incomp.sec	4194(62.97%)	5294(87.81%)	1
Compl.sec,			
Higher	2467(37.04%)	735(12.19%)	-1.ref
Delivery's place			
Hospital	3568(53.57%)	2119(35.14%)	1
Other	3093(46.43%)	3936(65.28%)	-1.ref
Birth order			
First to third	5632(71.5%)	3081(51.1%)	1
Above	1029(28.5%)	2948(48.90%)	-1.ref
Birth interval			
Less than 24	3093(46.43%)	1124(18.64%)	-1.ref
Greater than 24	3568(53.57%)	4905(81.36%)	1
Pregnancy's treatment			
Yes	697(10.46%)	1001(16.60%)	1
No	5964(89.54%)	5028(83.40%)	-1.ref
Receive vaccination			
Yes	1737(25%)	1299 (21.5%)	1
No	56(0.8%)	2923(48.48%)	-1.ref
Missing	75%	30%	
Drinking water			
Controlled	5374(80.68%)	1899(32%)	1
Not controlled	1287(19.32%)	4096(67%)	-1.ref
Missing		1%	
Has radio			
Yes	5374(80.68%)	4466(74.08%)	1
No	1559(19.32%)	563(25.92%)	-1.ref
Has electricity			
Yes	6203(93.12%)	2715(45.03%)	1
No	458(6.88%)	3314(54.97%)	-1.ref
Toilet facility			
Own flush toile facility	1768(28%)	590(10%)	1
Other and no toilet facility	4511(71.8%)	5335(88.5%)	-1.ref
Missing	1%	1.5%	
Antenatal visit			
Yes	4181(63%)	2412(40%)	1
No	2342(35%)	1264(21%)	-1.ref
Missing	2%	29%	

Table 1.5: Factors analyzed in diseases and malnutrtion studies

under five years. Child's sex is male or female (reference category).

Exposure to mass media

The 2003 DHS collected information on the exposure of women to various mass media including TV, radio, and print (i.e., magazines and newspapers). However, we will concentrate here on the radio ownership. Radio ownership is used as a simple indicator of socio-economic status. Lack of radio (reference category). In Egypt, as noted in pervious surveys, television has the widest coverage of the three media (TV, radio, and print). Lack of various mass media might result in less exposure education messages about management of common childhood diseases, infant feeding practices, and importance of vaccination.

Household socio-economic characteristics

The DHS gathered information on housing characteristics such as electricity, source of water and type of toilet facilities.

Electricity availability

Electricity is important for families to have access to electronic assets such as TV and radio. About 50 percent of households in Nigeria have electricity, it is much more common in urban than in rural areas, while in Egypt overall, 93 percent of households have electricity and the differentials in availability by residence are small. Lack of electricity (reference category) could have many disadvantages, especially in family education.

Type of toilet facility

A health indication of the household is assessed through a socio-economic indicator as the type of facility which is recoded in three categories called "flush toilet," "traditional toilet," and "no toilet facility" (reference category), but in the later analysis the factor is recoded to two categories called "called "flush toilet" and "others" (reference category). Lack of sanitary facilities poses a serious public health problem (DHS 2003). The 2003 NDHS informed that only 15 percent of households have a flush toilet, while 57 percent use traditional pit toilets, and 25 percent have no facility. The 2003 EDHS reports that, in contrast, 80 percent of households in the urban governorates have a modern flush toilets compared to 8 percent among rural areas in Upper Egypt.

Source of drinking water

The report of UNICEF in 2002 indicates that more than half the world's population used water from a piped connection at home. Moreover, 92% of the urban population and 70% of rural papulation in developing countries use improved drinking water source. Use of improved water and sanitation has a lot of benefits: reduction of diseases (particularly diarrhea), avoided illness health-related costs; saving time associated with getting water, and that the sanitation facilities located closer to home.

As known presently there is a wide spectrum of waterborne disease such as cholera, trachoma, typhoid and paratyphoid. The most common disease is diarrhea, which can lead to morbidity and in many cases mortality. Source of drinking water is recorded with respect to the water's quality, where water in the residence or from public tap is assumed to have controlled quality. However, water from a public well, spring, river stream, pond, lake or rainwater is not controlled. Tanker's water is assumed to contain uncontrolled water because it is scarce and costly.

Birth order

In some of previous studies, the order of birth has been recoded into four categories assumed that a higher order births are associated with a high risk of mortality. In this work, birth order is recorded into three categories: first to third birth, and higher (ref.cat).

Place of delivery

DHS collected information on the place of delivery for all births during the five years preceding the survey. We distinguish between the mothers who are delivered at hospital and the mothers who are delivered somewhere else (reference category).

Treatment during the pregnancy

The earlier coming of mother for antenatal care in her pregnancy helping for the earlier diagnosis and treatment of infections, and gives an opportunity to prevent low birth weight and other conditions in the newborn (UNICEF, 2004). The covariate indicates whether the mother received any treatment during the pregnancy or not (reference category).

Vaccination

Increasing the proportion of children who are vaccinated against the major preventable diseases of childhood is a cornerstone of survival programs. The information from the DHS considers the prevalence and treatment of diarrhea and acute respiratory infections illnesses that are among the most common causes of childhood deaths in these developing countries. This covariate is assumed as a binary factor, indicating whether the child is vaccinated.

Antenatal visit

Antenatal visits are recommended during a woman's pregnancy to ensure proper care. We assumed, overall, the women who obtained antenatal care (i.e. they made one or more visits to a provider) and the women who did not obtain antenatal visit during her pregnancy (reference category).

Chapter 2

Bayesian Geoadditive Models

Abstract

Generalized Additive Models are methods and techniques developed and popularized by Hastie and Tibshirani (1990). We examine the generalized geoadditive model as an alternative to the common linear model in the context of analyzing childhood disease and childhood malnutrition in Egypt and Nigeria. Most applications are still based on generalized linear models, assuming that covariate effects can be modeled by a parametric linear predictor. In our application, however, the data contain detailed information on metrical and geographical covariates. Their effects are often highly nonlinear, and are difficult to assess with conventional parametric models. In this chapter, we propose generalized geoadditive models which can simultaneously incorporate usual linear effects as well as nonlinear effects of metrical and spatial covariates within a semiparametric Bayesian approach. Inference is fully Bayesian and uses recent Markov Chain Monte Carlo (MCMC) simulation techniques for drawing random samples from the posterior.

2.1 Introduction

2.1.1 Generalized linear models

A common way to build regression models extending the classical linear model for Gaussian responses to more general situations such as binary responses are generalized linear models, originally introduced by Nelder and Wedderburn (1972). For more overviews see Fahrmeir and Tutz (2001) or McCullagh and Nelder (1989). In these models the influence of covariates on a response variable y is assumed to satisfy following two assumptions:

Distributional assumption

Conditional on covariates x , the responses y are independent and the distribution

y_i belongs to a simple exponential family, i. e. its density can be written as

$$p(y_i|x_i) = exp\left\{\frac{[y_i\theta_i - b(\theta_i)]}{\phi} w_i + c(y_i, \theta_i, w_i)\right\}, \quad i = 1, .., n \tag{2.1}$$

where

θ_i is the natural parameter of the exponential family

ϕ is a scale or dispersion parameter common to all observations

w_i is a weight i, and $b(.)$ and $c(.)$ are functions depending on the specific exponential family.

Structural assumption:

The (conditional) expectation $E(y|x) = \mu$ is linked to the linear predictor

$$\eta_i = x_i'\beta, \tag{2.2}$$

via

$$\mu_i = h(\eta_i) or \quad \eta_i = g(\mu_i)$$

where

the design vector x_i usually includes the *grand mean*

h is a smooth, bijective response function,

g is the inverse of h called the link function and

β is a vector of unknown regression coefficients.

Both assumptions are connected by the fact that the mean of y is also determined by the distributional assumption and can be shown to be given as

$$\mu_i = E(y_i|x_i) = b'(\theta_i),$$

In addition, $var(y_i|x_i)$ is the variance of y_i in general which is depend on the linear predictor with $\frac{\phi\nu(\mu_i)}{w_i} = b''(\theta_i)$ being the variance function of the underlying exponential family.

$$\sigma^2(\mu_i) = var(y_i|x_i) = b''(\theta_i)/w_i$$

The distribution of y_i could be normal, possion and binary (binomial) or any other exponential family distribution.

2.1.2 Models for Continuous Responses

Normal distribution

The classical linear model can be subsumed into the context of generalized linear models by defining $h(h) = \mu$, i. e. the response function is simply the identity. For Gaussian distributed responses this also represents the natural link function. The variance function $\nu(\mu)$ is constant, while the scale parameter equals the variance of the error terms of the linear regression model.

2.1.3 Models for Binary and Binomial Responses

For binary responses $y \in (0, 1)$ the expectation is given by the probability $\pi = P(y = 1)$, which requires appropriate response functions to ensure

$\pi \in [0,1]$. Obviously, any cumulative distribution function satisfies this condition and different model formulations are obtained for different choices of the distribution function. In any case, the scale parameter is again fixed at $\phi = 1$.

Logit model

When choosing the natural link function

$$g(\pi) = log\frac{\pi}{1-\pi} = \eta,$$

the logit model is obtained, which corresponds to the logistic distribution function as response function:

$$h(\eta) = \frac{exp(\eta)}{1 - exp(\eta)} = \pi. \tag{2.3}$$

The logistic distribution function is symmetric and has somewhat heavier tails than the standard normal distribution function used in probit models. The logit model is most commonly used when analyzing binary data, especially in medical applications. The generalized linear model differs from the general linear model (of which multiple regression is a special case) in two major aspects:

Firstly, the distribution of the dependent or response variable can be (explicitly) non-normal, and does not have to be continuous, e.g., it can be possion.

Secondly, the prediction of the dependent variable is based on a linear combination of predictor variables, which are connected to the dependent variable via a link function.

Probit model

The logistic distribution function is replaced by the the standard normal distribution function in the probit model. Since the evaluation of the likelihood for the probit model is computationally more demanding and parameter estimates are not interpretable in terms of odds or odd ratios, the logit model is often preferred.

2.2 Bayesian Geoadditive Models

The assumption of a parametric linear predictor for assessing the influence of covariate effects on responses seems to be rigid and restrictive in our practical application situation and also in many real statistically complex situation since their forms cannot be predetermined *a priori*. Besides, practical experience has shown that metrical covariates often have nonlinear effects. We are facing one of the following problems:

- In our application, for the continuous covariates in the data set, the assumption of a strictly linear effect on the predictor may not be appropriate, i. e. some effects may be of unknown nonlinear form (such as child's age, mother's age and mother's BMI). These variables have a nonlinear effect on the response variables.
- Another difficulty is that we have a spatial covariate in our models.

  Hence, it is necessary to seek for a more flexible approach for estimating the metrical covariates by relaxing the parametric linear assumptions. This in turn allows data to know the true functional form of the metrical effects and this approach is referred to as nonparametric regression model. To specify a nonparametric regression model, an appropriate function that contains the unknown regression function needs to be chosen. This choice is usually motivated by smoothness properties, which the regression function can be assumed to possess. To overcome these difficulties, we replace the strictly linear predictor in 2.2 by a geoadditive predictor.

Observation model

Suppose that regression data consists of observations, (y_i, x_i, w_i), $i = 1, ..., n$ on a response y_i. The response variables in this application will used logit model, probit model in the case of childhood disease and Gaussian model in the case of childhood undernutrition. We have to distinguish between a vector $x_i = (x_{i1}, ..., x_{ip})$ which is not necessary to be just metrical covariates but may be also time scales or spatial covariates, and w_i =$(w_{i1}, ..., w_{ip})'$ of covariates, whose effect is modelled in the usual form. In this application, the metrical covariates include the child's age, mother's age at birth and mother's BMI, a spatial covariate, which including the district in which the most of child's disease and child's undernutrition may be considered. In this application w_i will include categorical variables which are coded in effect code such as such child'sex, educational level of mother,..,etc. (see table 1.3.3). Generalized additive and semiparametric models (Hastie and Tibshirani, 1990) assume that, given $x_i = (x_{i1}, ..., x_{ip})$, and w_i the distribution of y_i belongs to an exponential family, with mean $\mu_i = E(y_i|x_i)$ linked to an additive predictor η_i by an appropriate response function h. We assume a semiparamtric regression model with geoadditive predictors

$$\mu_i = h(\eta_i), \quad \eta^{geo} = f_1(x_{i1}) + ... + f_p(x_{ip}) + f_{spat} + w_i'\gamma \tag{2.4}$$

Here h is a known response function, and $f_1, ..., f_p$ are possibly nonlinear functions of metrical covariates and f_{spat} is the effect of the spatial covariate $s_i \in 1, ..., S$ labeling the districts in the two countries. Regression models with predictors as in 2.4 are sometimes referred to as geoadditive models. In a further step we may split up the spatial effect f_{spat} into a spatially correlated (structured) and uncorrelated spatial (unstructured) effect

$$f_{spat}(s_i) = f_{str}(s_i) + f_{unstr}(s_i) \tag{2.5}$$

One rationale is that a spatial effect is usually a surrogate of many unobserved influences, some of them may obey a strong spatial structure and others may be present only locally. By estimating a structured and an unstructured effect we attempt to separate these effects.

2.3 Prior Distributions

In Bayesian inference, the unknown functions f_j, the fixed effects parameters γ as well as the variance parameter σ^2 are considered as random variables and have to be supplemented by appropriate priors distribution.

2.3.1 The General Form of the Priors

Suppose that $f = (f(1), ..., f(n))'$ be the vector of corresponding function evaluations at observed values of x.

Then, the general form of the prior for f is

$$f|\tau^2 \propto exp(-\frac{1}{2\tau^2}f'Kf) \tag{2.6}$$

Where K is a penalty matrix that penalizes too abrupt jumps between neighboring parameters. In most cases K will be rank deficient, therefore the prior for f would be improper. This implies that $f|\tau^2$ follows a partially improper Gaussian prior $f|\tau^2 \sim N(0, \tau^2 K^-)$ where K^- is a generalized inverse of a band-diagonal precision or penalty matrix K.

In the frequentist approach the smoothing parameter is the equivalent with the variance parameter τ^2 which controls the trade off between flexibility and smoothness. In order to estimate the smoothness parameter f, a highly dispersed but proper hyperprior is assigned to τ^2. The proper prior for τ^2 is required to obtain a proper posterior for f (Hobert and Casella, 1996). We choose an inverse gamma distribution with hyperparamters a and b, i.e.

$$\tau^2 \sim IG(a, b).$$

A particular prior depends on the type of the covariates and on prior beliefs about smoothness of f.

Furthermore, a prior for a function f is defined by specifying a smoothness prior, and the hyperparameters a and b of the inverse gamma prior for τ^2.

A possible choice for a and b is very small $a = b$, for example $a = b = 0.0001$, leading to almost diffuse priors for the variance parameters. An alternative proposed, for example, in Besag et al.(1995) is $a = 1$ and small value for b, such as $b = 0.005$. The choice of such a highly dispersed but proper prior avoids problems arising with improper priors. Such problems are discussed in Hobert and Casella (1996) for linear mixed models.

2.3.2 Priors for Fixed Effects

For the parameter vector γ of fixed effects we choose a diffuse prior

$$\gamma_j \propto const, \text{ j=1,...,r.}$$

Another choice would be to work with a multivariate Gaussian distribution $\gamma \sim N(\gamma_0, \Sigma_{\gamma_0})$. In this application, the diffuse priors will be used for the fixed effects.

2.3.3 Priors for Metrical (Continuous) Effects

Several alternatives are available to specify the priors of the unknown (smooth) functions $f_j, j = 1, .., p$. These are basis function approaches with adaptive knot selection (e.g. Dension et al., 1998, Biller, 2000) and approaches based on smoothness priors. In addition, several alternatives have been recently proposed for specifying a smoothness prior for the effect f of metrical covariate x. These are random walk priors (Fahrmeir and Lang, 2001), Bayesian smoothing splines (Hastie and Tibshirani, 2000) and Bayesian P-splines (Lang and Brezger, 2005). Our focus in this work is on random walk and P-splines priors.

First and second order random walk

Let us consider the case of a metrical covariate x with *equally-spaced observations* $x_i, i = 1, ...m, m \leq n$. Then $x_{(1)} < ..... < x_{(m)}$ defines the ordered sequence of distinct covariate values. Here m denotes the number of different observations for x in the data set. A common approach in dynamic or state space models is to estimate one parameter $f(t)$ for each distinct $x(t), i.e.,$. Define, $f(t) =: f(x_{(t)})$ and let $f = (f(1), .., f(t), .., f(m))'$ denote the vector of function evaluation. Then a first order random walk prior for f is defined by

$$f(t) = f(t-1) + u(t) \tag{2.7}$$

A second order random walk is given by

$$f(t) = 2f(t-1) - f(t-2) + u(t), \tag{2.8}$$

$$u(t) \sim N(0; \tau^2)$$

with diffuse priors $f(1) \propto const$ and $f(2) \propto const$, for initial values, respectively. A first order random walk penalizes too abrupt jumps $f(t) - f(t-1)$ between successive states. While, a second order random walk penalizes deviations from the linear tread $2f(t-1) - f(t-2) + u(t)$. In addition, the variance τ^2 controls the degree of smoothness f.

$$f_t|f_{t-1}, \tau^2 \sim N(f_{t-1}, \tau^2) \tag{2.9}$$

Random walk priors may be equivalently defined in a more symmetric form by specifying the conditional distributions of function $f(t)$ given its left and right neighbors. That means, we can write the prior in (2.7 and 2.8) in general form as

$$f|\tau^2 exp\left(-\frac{1}{\tau^2}f'Kf\right) \tag{2.10}$$

Here the design matrix K is the penalty matrix that penalizes too abrupt jumps between neighboring parameters. More often, K is not full rank and this implies that $f|\tau^2$ follows a partially improper Gaussian prior

$$f|\tau^2 \sim N(0, \tau^2 K^-)$$

where K^- is a generalized inverse of the penalty matrix K.

For the case of nonequally spaced observations, random walk or autoregressive priors have to be modified to account for nonequal distances $\delta_t = x(t) - x(t-1)$ between observations.

Random walks of first order are now specified by

$$f(t) = f(t-1) + u(t), \tag{2.11}$$

$$u(t)\ N(0; \delta_t \tau^2),$$

i.e., by adjusting from τ^2 to $\delta_t(\tau^2)$.

Random walks second order are

$$f(t) = \left(1 + \frac{\delta_t}{\delta_{t-1}}\right) f(t-1) - (\frac{\delta_t}{\delta_{t-1}}) f(t-2) + u(t), \tag{2.12}$$

$$u \sim N(0; w_t \tau^2),$$

where w_t is an appropriate weight. Several possibilities are conceivable for weights. The simplest one is $w_t = \delta_t$ for the first order random walk, see Fahrmeir and Lang (2001a) for a discussion.

Bayesian P-splines

A closely related approach for metrical covariates is based on the P-splines approach, introduced by Eilers and Marx (1996). The basic assumption of this approach is that the unknown function f_j can be approximated by a spline of degree l with equally spaced knots $x_{min} = \xi_0 < \xi_1 < ... < \xi_{r-1} < \xi_r = x_{max}$ within the domain of x_j. The domain from x_{min} to x_{max} can be divided into n' equal intervals by $n'+1$ knots. Each intervals will be covered by $l+1$ B-splines of degree l. The total number of knots for construction of the B-splines will be $n' + 2l + 1$. The number of B-splines in the regression is $n = n' + l$. It is well known that such a spline can be written in terms of a linear combination of $M = r + l$ B-splines basis functions β_j, i.e

$$f_j(x_{ij}) = \Sigma_{p=1}^{M} \beta_j B_j(x).$$

The basis functions B_j are defined locally in the sense that they are nonzero only on a domain spanned by $2 + l$ knots. The $n \times M$ design matrix X_j for P-splines is more intricate than the case of random walk priors. Each row i contains the value of the B-spline basis functions evaluated at x_i, hence $X_j(i, p) = B_{jp}(x_{ij})$. In accordance with the properties of B-splines (see De Boor, 1978), each row X has $M + 1$ non-zero values. As for the number of knots, Eilers and Marx (1996) recommended the number of inner knots to range between 20 and 40 and introduced a penalization of the differences between regression coefficients of adjacent B-spline basis functions in order to generate a smoothing effect. In our analysis, we typically choose B-splines of degree =3 and 10 intervals, and second order random walk priors on the B-splines regression coefficients.

Spatial Covariates

Consider first that the spatial index $s \in \{1, .., S\}$ represents a location or site in connected geographical regions. It is assumed that neighboring sites that share boundaries are more homogenous than any other arbitrary sites. Therefore, for a valid prior definition a set of neighbors must be defined for each site s. Hence sites s and t are neighbors if they share a common boundary. Depending on the application, the spatial effect may be further split into a spatially correlated (structured) and an uncorrelated (unstructured) effect, i.e. $f_{spat} = f_{str} + f_{unstr}$. A rationale is that a spatial effect is usually a surrogate of many unobserved influential factors, some of them may obey a strong spatial structure while others may exist only locally. Besag, York and Mollie (1991) proposed a Markov random field prior for the correlated spatial effects f_{str}. The spatial smoothness prior of function evaluations $f_{str}(s)$ is

$$f_{str,s}|f_{str,t}, t \neq s, \tau^2 \sim N\left(\sum_{t \in \delta_s} \frac{f_{str,t}}{N_s}, \frac{\tau^2_{str}}{N_s}\right), \tag{2.13}$$

where N_s is the number of adjacent sites and $t \in \delta_s$ denotes, that site f_s is a neighbor of site f_t. Thus the (conditional) mean of f_s is an unweighted average of function evaluations of neighboring sites. Note that for spatial data conditioning is undirected since there is no natural ordering of different sites f_s as in the case for metrical covariates.

In a general form, (2.13) can be given by

$$f_{str,s}|f_{str,t}, t \neq s, \tau^2 \sim N\left(\sum_{t \in \delta_s} \frac{w_{st}}{w_{s+}} f_{str,t}, \frac{\tau^2_{str}}{w_{s+}}\right), \tag{2.14}$$

where w_{sj} are known equal weights and w_{s+} denotes the marginal sum of w_{st} over the missing subscript. Such a prior is called a Gaussian intrinsic autoregression. For more details, see Besag et al. (1991), Besag and Kooperberg (1995).

The design matrix X_{str} is a $n \times S$ incidence matrix whose entry in the i-th row and s-th column is equal to one if observation i has been observed at location s and zero otherwise.

For the uncorrelated effect, we assume i.i.d. Gaussian random effects, i.e.

$$f_{unstr}(s) \sim N(0, \tau^2_{unstr}) \quad s = 1, .., S$$

Formally, the priors for f_{str} and f_{unstr} can both be brought into the form (2.10). For f_{str}, the elements of K given by

$$k_{ss} = w_s +$$

and

$$k_{st} = \begin{cases} w_{st} = -1 & \text{where} \quad t \in \delta_s \\ 0 & \text{otherwise} \end{cases}$$

For f_{unstr}, we may set $K = I$.

Furthermore, the inverse Gamma priors are assumed for $\tau^2_{str}[IG(a_{str}, b_{str})]$ and $\tau^2_{unstr}[IG(a_{unstr}, b_{unstr})]$.

2.4 MCMC Inference

We use Markov Chain Monte Carlo (MCMC) simulations to draw samples from the posterior. Statistical inference is done by means of Markov chain Monte Carlo techniques in a full Bayesian setting. We restrict the presentation to models with predictor 2.4. Full Bayesian inference is based on the entire posterior distribution.

$$p(\beta, \tau^2, \gamma|y) \propto p(y|\beta, \tau^2, \gamma)p(\beta, \tau^2, \gamma), \tag{2.15}$$

where $\beta = (\beta_1, .., \beta_p)$ and $\tau^2 = \tau^2_1, .., \tau^2_p$ denote parameter vectors for function evaluations and variance. Then, under usual conditional independence assumptions, the posterior is given by:

$$p(\beta, \tau^2, \gamma|y) \propto \prod_{i=1}^{n} L_i(y_i; \eta_i) \prod_{j=1}^{p} \{p(\beta_j|\tau^2_j)p(\tau^2_j)\} \prod_{k=1}^{r} p(\gamma_k)p(\sigma^2) \tag{2.16}$$

Only for Gaussian responses, the full conditional distributions for unknown functions $\beta_j, j = 1, .., p$, and fixed effects parameters γ are Gaussian and for variance components $\tau_j, j = 1, .., p$ and σ^2 the full conditionals are inverse gama distributions.

$$p(\beta|.) \propto \textstyle\prod_{i=1}^{n} L_i(y_i; \eta_i)p(\beta_j|\tau_j)$$

$$p(\gamma|.) \textstyle\prod_{i=1}^{n} L_i(y_i; p(\eta_i)p(\gamma)$$

$$p(\tau^2|.) = p(f|\tau^2)p(\tau^2)$$

$$p(\sigma^2|.) = \textstyle\prod_{k=1}^{r} p(\gamma_k)$$

Bayesian inference via MCMC is based on updating full conditionals of single parameters or blocks of parameters, given the rest and the data. For

Gaussian models, Gibbs sampling with so-called multimove steps can be applied. For non-Gaussian responses Gibbs sampling is no longer feasible and Metropolis Hastings algorithms are needed. More details can be found in Rue (2001) or Fahrmeir and Lang (2001a). For the predictor 2.4, let α denote the vector of all unknown parameters in the model. Then, under usual conditional assumptions, the predictor is given by

$$p(\alpha|y) \propto \Pi_{i=1}^{n} L_i(y_i, \eta_i) \Pi_{j=1}^{p} \{p(\beta_j|\tau_j^2) p(\tau_j^2)\} p(f_{str}|\tau_{str}^2) p(f_{unstr}|\tau_{unstr}^2) \Pi_{j=1}^{r} p(\gamma_j) p(\sigma^2),$$

where $\beta_j, j = 1, .., p$, are the vectors of regression coefficients corresponding to the functions f_j. The full conditionals f_{str}, f_{unstr} and fixed effects parameters γ are multivariate Gaussian in the case for Gaussian response variables. While the full conditionals for the variance components $\tau^2, j = 1, .., p$, str, unstr and σ^2 are inverse gamma distributions. More details can be found in Rue (2001), Fahrmeir and Lang (2001b), Lang and Brezger (2000a), and Kandala, et.al.(2001b). The estimation of models in this thesis is based on different sampling schemes depending on the distribution of the response. Two types of responses are included in this thesis, namely binary responses and Gaussian responses.

Gaussian Response

For the Gaussian response variable, the full conditionals for fixed effects and non-linear effects are multivariate Gaussian. For the variance parameters, all full conditionals are inverse Gamma distribution. Straight forward calculations show that precision matrices for nonlinear effects are band matrices. For a one dimensional P-spline the bandwidth of precision matrix is the maximum between the degree of the spline and the order of the random walk. The cholesky decomposition is mostly used for fast efficient matrix operation of band matrices. More details and description on the sampling scheme for Gaussian responses can be found in Lang and Brezger, 2001 and Rue, 2001.

Non-Gaussian Responses

Here, we now turn the attention to general responses from an exponential family. In this case the full conditionals are no longer Gaussian. For fixed effects and i.i.d. random effects we use a slightly modified version of the iteratively weighted least squares proposal suggested by Gamerman (1997), see also Brezger and Lang (2006), CSDA. In addition, Fahrmeir and Lang (2001a) propose a MH-algorithm for updating unknown regression parameters based on conditional prior proposals. For updating, only likelihood is required but no approximations of characteristics of the posterior (e.g. the mode).

Chapter 3

Modelling of Child Diseases in Egypt and Nigeria

Abstract

Our case study is based on the 2003 Demographic and Health Survey for Egypt (EDHS) and Nigeria (NDHS). It provided data on the prevalence and treatment of common childhood diseases such as diarrhea, cough and fever, which are seen as symptoms or indicators of children's health status, causing increased morbidity and mortality. The causes of childhood illnesses are multiple. Theses causes are associated with a number of risk factors, including inadequate antenatal care, lack of or inadequate vaccination, high birth order, and malnutrition. The main focus of this chapter is to analyze the effects of these different types of covariates on the response variables diarrhea, fever, and cough, using data from the 2003 DHS Demographic and Health surveys (DHS) from Egypt and Nigeria. We started our analysis using a large number of factors which could affect the health of children in both countries as a first step. Based on the results of the first step, we then excluded some factors which have slight effects on the childhood diseases as a second step and compare the results. A Bayesian geoadditive model for binary response variable is used in this application based on Fahrmeir and Lang (2001).

3.1 Introduction

In this application, we concentrate on flexible modelling of effects of metrical covariates, categorical covariates, and spatial covariates on the response variables (diarrhea, fever, and cough). The analyses for the childhood disease in Egypt and Nigeria are based on the data from the 2003 Demographic and Health survey (DHS). One of the main objectives of DHS is to provide an up-to-date information on childhood disease. This intends to assist policy makers and administrators in evaluating and designing programs and improve planning for future interventions in these areas, which in turn should reduce childhood morbidity and childhood mortality as well. We use the geoadditive logit models for the binary response variables (had diseases/no) in this chapter. Accordingly, we began the investigation with a large number of covariates including a large set of bio-demographic and socio-economic variables, including covariates such as preceding birth interval, current working status of mother, place of delivery, mother's educational attainment, whether the mother received injections during pregnancy or not and whether the mother attended antenatal clinic or not. Other relevant factors included such as mother's age at birth, availability of any toilet facility, source of drinkable water, locality of residence and region of residence. At the end, it turned out that many of them were not significant. The categorical covariates were transformed into effect coding. The metrical covariates are modelled by second order random walk priors. All computations have been carried out with BayesX-version 1.4 (Brezger, Kneib and Lang 2005).

3.2 Bayesian Models

In a first explanatory attempt, we fitted the data sets using a Bayesian linear model to model the effects of the covariates that clearly have linear effects on the child's disease. Next, we used flexible methods to model the metrical covariates which have nonlinear effects on the child's disease such as child's age, mother's age, and BMI of mother. Finally, we extended the model by including spatial determinants of child's disease and allocated these spatial effects to structured and unstructured (random) components.

3.2.1 Semiparametric Bayesian Regression Models

We estimate separate models for each disease in each country with predictor

$$\eta = f(x_{i1}) + ... + f_p(x_{ip}) + f_{spat}(s) + u_i'\gamma, \tag{3.1}$$

where $x_i = (x_{i1}, ..., x_{ip})'$ is a vector of covariates whose its influence assumed to be possibly nonlinear, and categorical covariates $u_i = (u_{i1}, ..., u_{iq})'$ with usual linear effects on the predictor. The functions $f_1, ..., f_p$ as well as the regression parameters γ are unknown and have to be estimated from the data. Moreover, $f_{spat}(s)$ is a spatial covariate which gives information about the location a particular observation pertains to. In a further step we split the spatial effect f_{spat} into correlated (f_{str}) and uncorrelated effect (f_{unstr}).

Therefore, we will use generalized geoadditive logistic models for binary response variables and the main focus of this stage is to analyze effects of these different types of covariates on the response variables diarrhea, fever, and cough.

The models which are obtained and discussed in this work would be validated by the DIC and deviance, which decrease from models with covariates of high explanatory value.

Deviance Information Criterion (DIC)

The classical approach to model comparison involves a trade-off between how well the model fits the data and the level of complexity. Spiegelhalter et al.(2002) devised a selection criterion which was based on Bayesian measures of model complexity and how good a fit the model is for the data. The measure of complexity which we adopted in this work is suggested by Spiegelhalter et al. (2002). A complexity measure pD is suggested by using an information theoretic argument to get more effective number of parameters in a model, as the difference between the posterior mean of the deviance and the deviance at the posterior estimates of the parameters of interest.

pD is assumed to be an approximate trace of the product of Fisher's information and the posterior covariance matrix. It could be obtained through a Markov Chain Mont Carlo analysis. In the case of normal models, pD corresponds to the trace of 'hat' matrix projection observations onto fitted values. In an exponential family model, $\bar{D}$ which calls for a posterior mean deviance, can be taken as a measure of fit. Assume that $f(y)$ is a fully specified standardizing term, then

$$pD = \bar{D}(\bar{\theta}) - D(\bar{\theta}), \tag{3.2}$$

where $D(\theta) = -2logp(y \mid \theta) + 2logf(y)$, is a Bayesian deviance.

A Deviance Information Criteria (DIC), which could be used for model comparison, is computed by adding the fit $\bar{D}$ to a complexity pD.

DIC is defined as a "Plug in" estimate of fit plus twice the effective number of parameter, as follows:

$$DIC = D(\bar{\theta}) + 2pD = \bar{D} + pD, \tag{3.3}$$

where the posterior mean of the deviance $\bar{D}(\theta)$ is penalized by the effective number of model parameters pD. See Spiegelhalter et al.(2002) for more details.

3.3 Statistical Inference

Bayesian geoadditive logit models were fitted to the three types of diseases of this data set.

The results for the following logit models presented in this application are selected from a larger hierarchy of models. For model choice and comparison we routinely use the Deviance Information Criterion (DIC) as mentioned above which was developed by Spiegelhalter et al. (2002). We need to point out that many models were utilized in the pre-analysis but only results of the selected model are discussed in this chapter. The following covariates were considered in the analysis to study childhood disease in Egypt and Nigeria.

Metrical covariates

Chage: Child's age in months.

BMI: Mother's body mass index.

Mageb: Mother's age at birth.

Categorical covariates (in effect coding)

male: Child's sex : male or female (reference category).

educ: Mother's educational attainment: complete primary education and incomplete secondary school "educp," complete secondary school and higher "educh," "no education," and "incomplete primary education" (reference category).

trepr: Whether mother had treatment during pregnancy: yes or no (reference).

anvis: Whether mother had antenatal care: yes or no (reference).

water: Source of drinking water: controlled water or no (reference category).

toilet: Has flush toilet at household "toiletf," has traditional toilet at household "toiletd," other type of toilet or no toilet (reference category).

urban: Locality where respondent lives : urban or rural (reference category).

radio: Has a radio at household: yes or no (reference category).

elect: Has electricity : yes or no (reference category).

work: Mother's current working status: working or not (reference).

bord: Birth order: first to third "bord," above third (reference category).

hosp: Place of delivery: hospital "hosp," other places (reference category).

vac: Receive vaccination: yes "vac" or no (reference category).

inv: Birth interval: More than 24 months "inv," less than 24 months (reference category).

Spatial covariate

reg: Governorates or regions where the respondent resides.

The responses y_j, $j = 1$ (diarrhea), 2 (fever), 3 (cough) are coded in this application as follows;

$$y_i = \begin{cases} 1 & : \text{if child had disease 2 weeks prior to the survey} \\ 0 & \text{if not} \end{cases} \tag{3.4}$$

The predictors of the models assumed in this section are as follows:

M1: This includes all categorical bio-demographic and socio-economic factors mentioned above

$$M1 : \eta_{ij} = \beta_0 + z_i' \gamma_i \tag{3.5}$$

M2: Adds the nonlinear effects to Model 1 and the vector z is reduced by omitting covariate ever had vaccination (vac)

$$M2 : \eta_{ij} = \beta_0 + f_j(chage) + f_j(BMI) + f_j(mageb) + u_i' \gamma_i \tag{3.6}$$

M3: The district-specific effects were added to the significant covariates in model 2

$$M3 : \eta_{ij} = \beta_0 + f_j(chage) + f_j(BMI) + f_j(mageb) + f_{str}(reg) + f_{unstr}(reg) + w_i' \gamma_i \tag{3.7}$$

In these models, β_0 is a constant term and the covariate vector z in model M1 contains all categorical bio-demographic and health factors. As for model M2, the vector z is reduced to the vector u by omitting factor of ever had vaccination. On the other hand, the nonlinear effects of the metrical covariates were included in M2. Model 3 contains the covariates which have significant effects on the disease based on the result of model 2. Furthermore, M3 includes the spatial effect f_{spat}. Moreover, we split the spatial effect f_{spat} into correlated (f_{str}) and uncorrelated effects (f_{unstr}) in M3.

3.4 Results

We began the investigation including a large set of bio-demographic and socio-economic variable, including covariates such as preceding birth interval, current working status of mother, place of delivery, mother's educational attainment, if the child has ever been vaccinated or not, whether the mother received antenatal care during pregnancy or not, and some other factors like availability of any form of toilet facility, source of drinkable water and place of residence, etc.

Starting with very simple models, we increase complexity to show what can be gained by more sophisticated approaches, and then we end up with the analysis using models that included the significant effects of categorical covariates as well as the nonlinear effects and the spatial effects. Furthermore, these models should be best in terms of DIC too.

The results show that, model M1 displays all of the fixed covariates. However, there are only 1478 (78% missing values) out of 6661 observations used in M1 for the data set of Egypt. For Nigeria there were 2650 (60% missing values) out of 6029 observations used in M1. The reason for that is the highly percentages of missing values which are included in variable "ever had vaccination". For this reason, it is not included in model M2. Furthermore, the results of model M1 for both countries, indicate that many of the covariates have nonsignificant effects on the three types of diseases with exceptions for some covariates. On the other hand, there are considerably high missing values including, those for variable "ever a mother obtained clinical visit" in the Nigeria's data set. However, we keep this variable in the analysis because it is assumed to play a role in childhood disease by previous health literature. Table 5.1 shows that M1 is best in terms of DIC, though we cannot compare its DIC with the DIC of model M1 and M2. The reason is the sample size used in M1 is very small compared to the sample size used in M2 and M3, which results a small value of DIC for M1. Therefore, we focus on model M2 and M3 in our discussion of this application. Moreover, M3 is the best in term of DIC compared to M2.

Regarding model M2, we excluded the covariates of "ever have been vacci-

Model	Deviance	pD	DIC
	Diarrhea		
M1(Egypt)	1282.92	16.62	1316.16
M2(Egypt)	5421.65	38.66	5498.98
M3(Egypt)	5300.038	44.22	5388.48
M1(Nigeria)	2629.19	16.71	2662.62
M2(Nigeria)	3130.01	30.77	3191.55
M3(Nigeria)	2949.1	45.14	3040.04
	Fever		
M1(Egypt)	1754.40	16.93	1788.26
M2(Egypt)	6974.67	38.06	7050.81
M3(Egypt)	6919.65	38.34	6996.33
M1(Nigeria)	3945.21	31.97	4009.16
M2(Nigeria)	5596.01	51.77	5699.55
M3(Nigeria)	5596.15	50.52	5697.21
	Cough		
M1(Egypt)	1616.5	16.47	1649.47
M2(Egypt)	6429.53	37.97	6505.47
M3(Egypt)	6344.32	35.45	6415.22
M1(Nigeria)	2902	16.84	2936.04
M2(Nigeria)	3587.88	33.65	3655.19
M3(Nigeria)	3417.712	49.839	3517.39

Table 3.1: Summary of the Deviance Information Criterion (DIC) for models M1 to M3 for both countries

nated" from the vector of categorical covariates, and the nonlinear effects of child's age, mother's BMI, and mother's age at birth were added.

In model M3, only the factors that have significant effects (or at borderline) on the disease by model M2 were taken into account, besides the spatial effects, which are split into structured effects and unstructured effects.

Diarrhea

The results for Egypt show that most of the covariates have a significant effect on the disease of diarrhea (tables 3.2 through 3.4). The results indicate that a higher risk of diarrhea is associated with; males, the birth order of the child, in addition to mothers who had primary school at most, their attendance for clinical care and in those who had received treatment during their pregnancy. In contrast, the lower level of diarrhea is associated with children living in urban areas, whose mothers had higher education and have a radio in the household. For Nigeria, male children, children who have been vaccinated, children from mothers who had treatment during pregnancy are at a higher risk of diarrhea (table 3.6). In contrast children from mothers who made antenatal visits during pregnancy are at a lower risk of diarrhea. The results of models M2 and M3 make apparent that the effects of most of the significant covariates remain quite stable when including and excluding some other covariates. The results of M2 (table 3.6) indicate that children whose mothers had primary school at most and having access to radio at home are at a lower risk of diarrhea, whilst the traditional toilet and birth order (first-third born) are associated with a higher risk of diarrhea. In addition, for model M3 (table 3.6), it turned out that only mothers who had primary school at most have a significant effect on diarrhea disease in Nigeria.

Figures 3.1 and 3.2 display the posterior mean of nonlinear effects of child's age, mother's BMI and mother's age at birth for model 2 (left panel) and model 3 (right panel) for Egypt and Nigeria, respectively. The nonlinear effects of metrical covariates are based on second order random walk priors, but in the next chapters we have used P-spline priors. Results for Egypt show that the impact of a child's age on the diarrhea disease for model M2 and model M3 are very similar (figure 3.1). Shown are the posterior means and the pointwise credible intervals. Figures include pointwise 80% and 95% credible intervals. As suggested by the 2003 EDHS (Demographic and Health Survey of Egypt, 2003), we are able to see a continuous worsening of the diarrhea disease during the first 10-11 months of age, and maybe even during the first 24 months of age. The deterioration sets in right after birth and continuous until about 11 months of age. As shown in table 1.1, the worsening is associated with children at around 6-11 months of age, followed by 12-35 months of age. The figures support these suggestions, showing a high risk during the first 11 months of age and after that, the impact of age decreases. For Nigeria, the effect of child's age seems to be high in the first 10-12 months of age and tends to stabilized at a high level of risk until 20 months of age, declining thereafter (figure 3.2).

Results for Egypt show that the effect of BMI (second panel from the top) is very slight for mothers with a BMI less than 27, and a higher risk exists for mothers who have a BMI between 28 and 35, where a blip appears.

Furthermore, after a BMI of 35 or 36, the influence is going to be negligible. Moreover, there is no difference between the patterns of model 2 and model 3. For Nigeria (second panel of figure 3.2), the effect of mother's BMI is significant until the BMI of 30 and declines thereafter. This means that the effect of mothers who have a BMI between 15-30 is significant for diarrhea disease in Nigeria.

The bottom panels of figure 3.1 show the impact of mother's age on diarrhea in Egypt. The largest effect is found for younger mothers (less than 25 years old), and the effect declines thereafter. Furthermore, it shows that the effect of mothers younger than 20 is considerably higher compared to that of mothers in middle age (25-35). As for Nigeria, the effect of young mothers (less than 20 years) is relatively high compared to their counterparts (figure 3.2).

With regard to spatial effects, the geographical pattern of the regions in the right panel of figure 3.3 depicts the estimated posterior mean of the structured random effects on the diarrhea disease in Egypt. Obviously there exists a district-specific geographical variation in the level of the disease in Egypt based on the 2003 EDHS. It is revealed that significant high rates of illness are associated with the Upper Egypt area (Minya, Amarna, Luxor, Esna, Edfu, Aswan, Sinai, Port Said, Suez Canal, Damietta). An Upper Egypt area implies a relatively high risk of having diarrhea and knowing the characteristics of the region, the result is not a surprise. On the other hand, the lower Egypt area (essentially the region of Nile Delta such as Cairo, Alexandria) is associated with significantly lower rate of illness. The left panel of figure 3.3 shows the unstructured effects on diarrhea disease in Egypt. It is similar to the structured effects. However the gray area indicates that there are no observation found in this area. This not only for the diarrhea disease, but also for fever and cough as can be shown later. This means there are no children living in this area (desert area) as the data of Egypt suggests.

Figure 3.2 shows colored maps of the structured random effects on the diarrhea disease in Nigeria (left panel) and its corresponding of the unstructured effects (right panel). Figure 3.2 shows that most of children from south-

Variable	Mean	S.dv	2.5%	median	97.5%
const	-2.95*	0.526	-3.917	-2.947	-1.96
urban	−0.280*	0.090	-0.454	-0.283	-0.092
male	0.112	0.069	-0.018	0.111	0.254
educp	-0.010	0.143	-0.294	-0.0148	0.268
educh	-0.153	0.110	-0.366	-0.152	0.061
work	0.133	0.179	-0.221	0.137	0.484
toiletd	0.021	0.372	-0.741	0.018	0.759
toiletf	-0.042	0.272	-0.559	-0.041	0.488
radio	-0.050	0.093	-0.225	-0.053	0.143
elect	0.275	0.332	-0.293	0.263	0.931
anvis	0.179*	0.089	0.0003	0.178	0.342
inv	-0.140	0.097	-0.317	-0.144	0.048
bord	0.084	0.083	-0.081	0.085	0.243
trepr	0.025	0.120	-0.232	0.032	0.252
vac	0.796*	0.390	0.106	0.765	1.702
hosp	0.535*	0.1600	0.229	0.532	0.858
water	-0.111	0.1053	-0.320	-0.110	0.082

Table 3.2: Fixed effects (M1) on diarrhea for Egypt.

eastern and northern parts of the country are highly affected by diarrhea. Furthermore, these regions have a highly significant effect on infancy deaths (Adebayo, 2002). On the other hand, there are some regions that have negative significant effects or have non-significant effects on the diarrhea disease.

The left panel of figure 3.2 shows the unstructured effect. It is significant and similar to the structured effect. The unstructured effect suggests similar variation in the level of diarrhea disease as in the structured effects.

Fever

The results for the estimated categorical parameters indicate that the prevalence of fever is higher among male children from mothers who have antenatal care during pregnancy and currently work, as the results of Egypt suggest. The households having access to radio are associated with a lower risk of fever (tables 3.9 through 3.10). On the other hand, the lower level of fever is associated with mothers having completed higher education, but children who have been vaccinated are not associated with a lower risk of fever as the results of model M1 suggest (table 3.8). Results of Nigeria (tables 3.12 and 3.13), indicate that the prevalence of fever is lower among

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.72^{*}	0.268	-2.25	-1.72	-1.22
urban	-0.131^{*}	0.039	-0.217	-0.13	-0.052
male	0.106^{*}	0.033	0.037	0.109	0.167
educp	0.176^{*}	0.063	0.049	0.176	0.306
educh	-0.127^{*}	0.057	-0.239	-0.127	-0.016
work	0.075	0.092	-0.103	0.075	0.252
toiletd	0.052	0.161	-0.260	0.049	0.346
toiletf	-0.117	0.112	-0.337	-0.122	0.087
radio	-0.117^{*}	0.047	-0.214	-0.115	-0.027
elect	0.066	0.163	-0.222	0.067	0.377
anvis	0.120^{*}	0.039	0.049	0.117	0.204
inv	0.013	0.048	-0.076	0.011	0.111
bord	0.116^{*}	0.040	0.036	0.116	0.196
trepr	0.114^{*}	0.050	0.0142	0.113	0.214
hosp	0.022	0.07	-0.134	0.020	0.166
water	-0.012	0.05	-0.114	-0.010	0.080

Table 3.3: Fixed effects (M2) on diarrhea for Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.919^{*}	0.237	-2.418	-1.906	-1.476
urban	-0.117^{*}	0.042	-0.199	-0.116	-0.037
male	0.109^{*}	0.035	0.0416	0.108	0.186
educp	0.158^{*}	0.064	0.030	0.158	0.281
educh	-0.123^{*}	0.059	-0.247	-0.124	-0.005
radio	-0.093^{*}	0.043	-0.182	-0.090	-0.008
anvis	0.137^{*}	0.0400	0.065	0.137	0.215
bord	0.132^{*}	0.040	0.056	0.131	0.212
trepr	0.102^{*}	0.050	0.0004	0.101	0.197

Table 3.4: Fixed effects (M3) on diarrhea for Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.321^*	0.148	-1.609	-1.311	-1.017
urban	0.029	0.0602	-0.081	0.029	0.148
male	0.112^*	0.047	0.017	0.114	0.204
educp	-0.188	0.103	-0.405	-0.191	0.027
educh	-0.283	0.153	-0.581	-0.283	0.003
work	-0.086	0.102	-0.296	-0.090	0.118
toiletd	0.170	0.093	-0.0083	0.167	0.358
toiletf	0.002	0.170	-0.334	0.002	0.332
radio	-0.149^*	0.054	-0.262	-0.149	-0.041
elect	-0.098	0.058	-0.212	-0.099	0.013
anvis	-0.142^*	0.053	-0.257	-0.142	-0.040
inv	-0.053	0.062	-0.175	-0.052	0.073
bord	0.029	0.051	-0.073	0.031	0.133
terpr	0.153^*	0.066	0.022	0.155	0.292
vac	0.144^*	0.052	0.035	0.146	0.244
hosp	-0.216	0.136	-0.483	-0.215	0.048
water	-0.031	0.069	-0.167	-0.031	0.108

Table 3.5: Fixed effects (M1) on diarrhea for Nigeria.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-2.233^*	0.313	-2.876	-2.229	-1.657
urban	0.003	0.057	-0.111	0.004	0.120
male	0.086	0.044	-0.004	0.086	0.168
educp	-0.184^*	0.082	-0.346	-0.184	-0.012
educh	-0.072	0.117	-0.326	-0.074	0.146
work	0.0023	0.097	-0.202	-0.001	0.193
toiletd	0.202^*	0.073	0.062	0.201	0.348
toiletf	-0.114	0.131	-0.396	-0.103	0.148
radio	-0.128^*	0.052	-0.230	-0.127	-0.0246
elect	-0.072	0.055	-0.178	-0.073	0.036
anvis	-0.105	0.057	-0.221	-0.105	0.006
inv	-0.018	0.062	-0.138	-0.019	0.107
bord	0.021^*	0.046	-0.071	0.022	0.106
terpr	0.139^*	0.060	0.022	0.140	0.258
hosp	-0.128	0.118	-0.360	-0.129	0.095
water	-0.012	0.065	-0.142	-0.013	0.108

Table 3.6: Fixed effects (M2) on diarrhea for Nigeria.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-2.357^{*}	0.315	-3.027	-2.340	-1.804
educp	-0.136^{*}	0.061	-0.252	-0.135	-0.016
toiletd	0.060	0.053	-0.0462	0.0621	0.168
radio	-0.081	0.053	-0.181	-0.082	0.020
anvis	-0.0917	0.057	-0.209	-0.0919	0.022
bord	0.041	0.051	-0.059	0.043	0.141
trepr	0.094	0.055	-0.205	-0.096	0.0183

Table 3.7: Fixed effects (M3) on diarrhea for Nigeria.

mothers who delivered their children in a hospital or have a flush toilet in the household. Lower risk is also associated with mothers who had no more than primary education (table 3.12). However, this variable is not significant any more in model M3. In addition, the traditional toilet and treatment during pregnancy are associated with a higher risk of fever as the results of Nigeria indicate.

Figures 3.5 and 3.6 show the nonlinear effects of child's age, BMI and mother's age on fever for models M2 and M3 in Egypt and Nigeria, respectively. The patterns of child's age (top of figures 3.5 and 3.6) show that the children face a high risk of suffering from fever during the first 12 months of their life, and then it declines slowly thereafter in Egypt. However, in Nigeria the age effect is still high until 25-26 months of age. The effect of a BMI on fever is slight in both countries. However, in Egypt there is a comparably high effect for mothers with BMI between 30 and 35. This also was observed in the results of diarrhea.

The nonlinear effect of mother's age on fever (bottom panels of figures 3.5 and 3.6) displays that young mothers (less than 20) have a significant effect on fever compared to mothers who are in the age group of 25-35. Furthermore, the effect of mother's age is similar to diarrhea's and cough's for both countries.

Figure 3.5 reveals that significant high fever illness rates are associated with the following governorates; Suez, El arish, Ismalia, and the southwest region Sinia (see the discussion section 3.5). The spatial unstructured effect is mostly similar to the structured effect with the exception for the desert area, which has no information.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-2.122^*	0.510	-3.204	-2.093	-1.174
urban	-0.099	0.068	-0.244	-0.097	0.039
male	0.0011	0.059	-0.120	0.002	0.120
educp	0.179	0.113	-0.040	0.176	0.410
educh	-0.188^*	0.095	-0.367	-0.187	-0.003
work	0.378^*	0.129	0.136	0.378	0.623
toiletd	0.045	0.289	-0.490	0.042	0.619
toiletf	-0.095	0.227	-0.546	-0.096	0.348
radio	-0.020	0.079	-0.181	-0.020	0.137
elect	-0.011	0.260	-0.494	-0.031	0.544
anvis	0.099	0.067	-0.037	0.101	0.232
inv	-0.045	0.083	-0.210	-0.045	0.122
bord	0.075	0.070	-0.060	0.075	0.211
trepr	0.049	0.099	-0.145	0.053	0.244
vac	1.138^*	0.390	0.463	1.088	2.013
hosp	2.565	1.696	-0.362	2.655	5.297
water	0.108	0.090	-0.083	0.112	0.282

Table 3.8: Fixed effects (M1) on Fever for Egypt.

For southeastern Nigeria, (figure 3.8) and through some regions in north, significant high rates of fever are observed (as the result of diarrhea suggest). In addition, there are significants high rates of disease shown in some districts in the southwest such as Zanfana and Kebbi, but the risk is not high compared to the southeastern districts. Non significant effects are noticed across some western regions and in some central regions as well. The spatial unstructured district effects for fever turn out to show a spatial variation in Nigeria.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.718^*	0.477	-1.658	-0.691	0.160
urban	0.018	0.034	-0.049	0.018	0.089
male	0.072^*	0.029	0.015	0.075	0.132
educp	0.014	0.054	-0.089	0.015	0.113
educh	-0.057	0.047	-0.154	-0.058	0.035
work	0.177^*	0.075	0.021	0.176	0.322
toiletd	0.181	0.144	-0.101	0.177	0.476
toiletf	-0.061	0.096	-0.254	-0.060	0.123
radio	-0.104^*	0.037	-0.176	-0.104	-0.0293
elect	-0.192	0.147	-0.475	-0.197	0.123
anvis	0.121^*	0.034	0.054	0.122	0.187
inv	0.047	0.042	-0.032	0.049	0.133
bord	0.006	0.035	-0.065	0.006	0.077
trepr	0.039	0.045	-0.043	0.037	0.130
hosp	0.045	0.0690	-0.088	0.0423	0.185
water	0.037	0.0455	-0.050	0.0368	0.125

Table 3.9: Fixed effects (M2) on Fever for Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.936^*	0.168	-1.299	-0.934	-0.604
male	0.073^*	0.029	0.015	0.074	0.129
work	0.155	0.078	-0.007	0.159	0.308
radio	-0.119^*	0.036	-0.192	-0.120	-0.048
anvis	0.126^*	0.031	0.064	0.125	0.186

Table 3.10: Fixed effects (M3) on Fever for Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.524^*	0.126	-0.770	-0.524	-0.286
urban	-0.069	0.050	-0.165	-0.078	0.029
male	0.034	0.0415	-0.049	0.035	0.118
educp	-0.097	0.0816	-0.254	-0.098	0.063
educh	0.0123	0.117	-0.232	0.0157	0.246
work	-0.0117	0.091	-0.206	-0.0131	0.166
toiletd	0.153^*	0.0714	0.015	0.153	0.289
toiletf	-0.137	0.131	-0.386	-0.134	0.123
radio	-0.044	0.0517	-0.147	-0.044	0.062
elect	-0.0678	0.0518	-0.169	-0.069	0.033
anvis	0.033	0.050	-0.065	0.0322	0.135
inv	0.028	0.058	-0.079	0.029	0.144
bord	0.030	0.0446	-0.058	0.031	0.118
trepr	0.105	0.057	-0.005	0.106	0.219
vac	0.089	0.047	-0.004	0.0916	0.184
hosp	-0.365^*	0.106	-0.582	-0.356	-0.170
water	0.049	0.0634	-0.073	0.046	0.170

Table 3.11: Fixed effects (M1) on Fever for Nigeria.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.899^*	0.229	-1.372	-0.887	-0.461
urban	-0.083	0.048	-0.180	-0.083	0.015
male	0.017	0.040	-0.060	0.015	0.099
educp	-0.156^*	0.067	-0.280	-0.156	-0.023
educh	0.084	0.0915	-0.102	0.079	0.260
work	-0.003	0.088	-0.180	-0.008	0.168
toiletd	0.223^*	0.062	0.100	0.224	0.342
toiletf	-0.266^*	0.105	-0.479	-0.262	-0.06
radio	-0.055	0.046	-0.146	-0.055	0.038
elect	-0.055	0.049	-0.152	-0.057	0.040
anvis	0.071	0.048	-0.015	0.068	0.176
inv	0.051	0.054	-0.054	0.049	0.156
bord	0.0016	0.038	-0.071	0.001	0.074
trepr	0.134^*	0.052	0.022	0.137	0.233
hosp	-0.315^*	0.097	-0.507	-0.314	-0.119
water	0.047	0.054	-0.062	0.049	0.150

Table 3.12: Fixed effects (M2) on Fever for Nigeria.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.051^*	0.187	-1.445	-1.052	-0.705
educp	-0.080	0.056	-0.188	-0.078	0.031
educh	-0.022	0.075	-0.179	-0.020	0.121
toiletd	0.116^*	0.052	0.022	0.114	0.223
toiletf	-0.214^*	0.087	-0.396	-0.216	-0.048
trepr	0.162^*	0.042	0.078	0.1608	0.245
hosp	-0.169^*	0.084	-0.349	-0.168	-0.0073

Table 3.13: Fixed effects (M3) on Fever for Nigeria.

Cough

Results of fixed effects are shown in tables 3.14-3.19 for Egypt and Nigeria respectively. The fixed parameters (tables 3.15 and 3.16) show that the prevalence of cough in Egypt is higher among male children, children from mothers who are currently working, had antenatal visits during pregnancy, and children from households with no access to radio. The households with traditional toilet, the vaccination status, and urban have positive significant effect on cough disease (M1 and M2), but they are not significant any more in model M3. However, the analysis of this data set indicates that source of water, place of residence, availability of electricity, type of toilet, treatment during pregnancy, child's place of delivery, education attainment, higher birth interval (> 24 months) and a birth order have either slight or nonsignificant effect on cough disease in Egypt.

For Nigeria, tables 3.17 to 3.19 display that a lower risk of cough is associated with access to controlled water and availability of electricity in households (M2 and M3). However, the treatment during pregnancy has a positive effect on cough. In addition, most of the other covariates have little or no influence on cough risk in Nigeria. Furthermore, source of water, availability of electricity are nonsignificant as observed by M3 (table 3.19).

The results of the non-linear effect of child's age (figures 3.9 and 3.10) suggest quite similar patterns for diarrhea and fever. The same is true for mother's age in both countries. However, the effect of BMI (second panel of figure 3.9) seems to be comparably higher for mothers with BMI $> 30 - 35$ in Egypt, but in Nigeria (second panel of figure 3.10) the effect of BMI is slight on cough disease.

There is a strong effect on cough risk gradient in some governorates of the Nile Delta in Egypt (figure 3.11). Whilst in Nigeria (figure 3.12), the south-eastern part of Nigeria, doing with some districts in northern regions part of country are associated with cough risk, as the results of fever disease also show. The unstructured spatial effects for cough turn out to be significant.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.908^*	0.452	-2.730	-1.92	-0.963
urban	0.025	0.072	-0.126	0.025	0.163
male	0.031	0.063	-0.087	0.030	0.148
educp	-0.036	0.125	-0.288	-0.0408	0.221
educh	0.054	0.101	-0.153	0.054	0.248
work	0.227	0.147	-0.050	0.229	0.531
toiletd	0.516	0.296	-0.0495	0.513	1.092
toiletf	-0.220	0.240	-0.688	-0.226	0.240
radio	-0.162^*	0.077	-0.315	-0.165	-0.001
elect	-0.058	0.270	-0.570	-0.048	0.490
anvis	0.035	0.076	-0.120	0.039	0.174
inv	-0.032	0.083	-0.198	-0.031	0.136
bord	0.057	0.068	-0.074	0.054	0.192
trepr	-0.005	0.106	-0.225	0.0003	0.197
vac	0.981^*	0.370	0.346	0.959	1.803
hosp	0.237	0.144	-0.045	0.234	0.523
water	0.036	0.095	-0.143	0.037	0.220

Table 3.14: Fixed effects (M1) on Cough for Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.05^*	0.215	-1.494	-1.059	-0.627
urban	0.077^*	0.036	0.008	0.076	0.155
male	0.069^*	0.031	0.006	0.070	0.131
educp	0.023	0.058	-0.088	0.026	0.134
educh	-0.069	0.055	-0.180	-0.069	0.037
work	0.242^*	0.081	0.081	0.244	0.392
toiletd	0.307^*	0.142	0.041	0.305	0.577
toiletf	-0.037	0.101	-0.235	-0.037	0.141
radio	-0.098^*	0.040	-0.179	-0.097	-0.019
elect	-0.069	0.153	-0.356	-0.074	0.249
anvis	0.124^*	0.036	0.052	0.128	0.193
inv	0.026	0.045	-0.0694	0.0272	0.112
bord	0.043	0.035	-0.022	0.041	0.116
trepr	0.020	0.050	-0.083	0.023	0.114
hosp	0.075	0.067	-0.059	0.077	0.197
water	-0.045	0.047	-0.133	-0.049	0.054

Table 3.15: Fixed effects (M2) on Cough for Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.988^*	0.169	-1.304	-0.985	-0.660
urban	0.054	0.036	-0.018	0.053	0.129
male	0.079^*	0.030	0.022	0.079	0.141
work	0.211^*	0.078	0.043	0.214	0.356
anvis	0.105^*	0.035	0.036	0.104	0.175
radio	-0.114^*	0.040	-0.195	-0.114	-0.034

Table 3.16: Fixed effects (M3) on Cough for Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.021^*	0.135	-1.280	-1.023	-0.746
urban	-0.007	0.058	-0.127	-0.008	0.103
male	-0.036	0.045	-0.127	-0.036	0.0504
educp	0.045	0.085	-0.123	0.044	0.216
educh	-0.139	0.129	-0.411	-0.133	0.107
work	-0.0139	0.093	-0.207	-0.014	0.153
toiletd	-0.045	0.077	-0.188	-0.048	0.104
toiletf	0.113	0.130	-0.154	0.113	0.365
radio	-0.0408	0.0529	-0.146	-0.040	0.06
elect	-0.094	0.055	-0.207	-0.098	0.009
anvis	0.047	0.053	-0.063	0.047	0.159
inv	0.024	0.061	-0.099	0.022	0.139
bord	0.004	0.048	-0.091	0.004	0.095
trepr	0.212^*	0.062	0.090	0.213	0.330
vac	0.124^*	0.052	0.023	0.127	0.226
hosp	-0.175	0.121	-0.417	-0.177	0.073
water	-0.219^*	0.065	-0.354	-0.218	-0.098

Table 3.17: Fixed effects (M1) on Cough for Nigeria.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.317^{*}	0.252	-1.859	-1.329	-0.810
urban	-0.040	0.050	-0.144	-0.042	0.065
male	-0.008	0.039	-0.087	-0.010	0.070
educp	0.014	0.068	-0.126	0.0156	0.144
educh	0.010	0.090	-0.172	0.011	0.197
work	0.101	0.087	-0.070	0.099	0.272
toiletd	0.034	0.061	-0.089	0.03	0.160
toiletf	0.054	0.103	-0.144	0.052	0.252
radio	-0.053	0.049	-0.151	-0.051	0.04
elect	-0.103^{*}	0.052	-0.210	-0.099	-0.004
anvis	0.075	0.050	-0.027	0.073	0.175
inv	0.051	0.060	-0.065	0.050	0.173
bord	0.031	0.043	-0.057	0.030	0.119
trepr	0.235^{*}	0.054	0.132	0.236	0.343
hosp	0.001	0.103	-0.200	0.003	0.204
water	-0.166^{*}	0.058	-0.276	-0.165	-0.049

Table 3.18: Fixed effects (M2) on Cough for Nigeria.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.139^{*}	0.191	-1.518	-1.135	-0.787
elect	0.011	0.040	-0.069	0.011	0.089
anvis	0.034	0.052	-0.0712	0.035	0.135
trepr	0.204^{*}	0.045	0.114	0.204	0.297
water	-0.0627	0.053	-0.159	-0.063	0.046

Table 3.19: Fixed effects (M3) on Cough for Nigeria.

3.5 Discussion and Conclusion

Fixed Effects

After investigating influential factors for the diseases, the fixed effects show the importance of child's sex, mother's employment attainment, antenatal visits during pregnancy, and availability of radio in households on the three types of diseases, besides the importance of mother's education, birth order, and place of residence on chronic diarrhea in Egypt. For Nigeria, the results indicate that the type of toilet (traditional toilet), availability of radio, birth order, mother's education (had primary education at most), and treatment during pregnancy are associated with diarrhea disease. The effects of having primary education, type of toilet, treatment during pregnancy, place of delivery are associated with fever disease, and treatment during pregnancy, availably of electricity and source of drinkable water have an influence on the disease of cough. The key results are quite plausible and consistent with the current literature.

Concerning child's gender, it is widely believed that the male disease is higher due to biological reasons, although, boys are markedly more likely than girls to be taken to treatment (EDHS, 2003). However, some studies show higher female morbidity, indicating gender discrimination. The results show that a child's gender is significant on the three types of diseases for model M2 and M3 in Egypt.

An interesting finding is that the level of maternal education is highly significant on diarrhea disease in Egypt. The results indicate that the children from mothers who have completed primary school are in a higher risk of diarrhea in Egypt, while they are in a lower risk of diarrhea in Nigeria. However, the children from mothers who are highly educated have a lower risk in Egypt.

We found that children from urban areas are at a lower risk of diarrhea disease in Egypt (tables 3.3 and 3.4 for M2 and M3, respectively), but this variable has a nonsignificant effect on the diseases in Nigeria. That is because the better services are more available and accessible in urban areas compared to rural areas. It reflects the role of public health policy to eliminate rural-urban disparities concerning health status, especially by improving sanitation infrastructure (and water as well) of rural areas, which could lead to the improvement of health status and reduce the rate of illness. On the other hand, because education is strongly correlated with welfare, the poor residents are in need of targeted efforts aimed at enhancing education opportunities (Poverty Reduction in Egypt Diagnosis and Strategy, 2002). Studies around the world have shown that more educated women, even in poor households, will typically have healthier children than less educated women (Child Health Diagnosis, 1995).

We also found that childhood morbidity was higher among mothers who had antenatal care during pregnancy (*anvis*) or had treatment during pregnancy (*trepr*). The reason for these unexpected results could be the following: The data set of Egypt indicates that there are few mothers had obtained antenatal visits frequently and 10% of mothers who had treatment during their pregnancy. On the other hand, the reason of visiting is not clear (whether it was related to their pregnancy or not). For Nigeria, this variable *anvis* is not significant as the data of Nigeria suggest. The treatment during pregnancy (*trepr*) has a positive significant effect on diarrhea and cough, because of same reason mentioned above.

It is interesting to note that the birth order is associated with a higher risk of diarrhea in both countries. However, it is not associated with fever or cough in both countries.

Having a radio in the household reduces the risk of the three types of diseases in Egypt, yet it is only associated with a lower risk of diarrhea in Nigeria. Ownerships of radio facilitates have a chance to get information allowing a more effective allocation of resources to produce the health of children (Kandala, 2002).

The results indicate a lower risk of cough in households having electricity in Nigeria, however this variable has no significant effect on the other types of diseases (fever and diarrhea) in Nigeria, and it has also no effect on the three types of diseases in Egypt.

With regard to current working status of the mother, the results of Egypt suggest a positive significant effect of this variable on fever and cough diseases (M2 and M3). In other words, children from mothers who work face a higher probability of getting fever and cough diseases. This stands in contrast to some previous studies which reported that mother's time, energy, knowledge, skills, her own health, along with the resources at her command, are critically important to the survival and healthy development of each of her children during the first months and even first few years of their lives. However, out-of home employment curtails the duration of full breastfeeding for many mothers. On the other hand, mothers with secondary education are employed in low-paid jobs and may not be able to afford adequate feeding of their babies.

Results for Egypt show that the type of toilet is not associated with the diseases of children, while in Nigeria the children from households using traditional toilets are in a relative higher risk of getting diarrhea, however the risk of fever is lower among the children from household using flush toilet.

The results show that the source of water has no significant effect on the diseases in Egypt. For Nigeria, the children from households using controlled water are at a lower risk of having cough disease. On the other hand, the direct relationship between access to water and the disease is indicated in previous studies, therefore, it is necessary to be concerned with how safe water reaches households. As reported in some earlier works, many households are not directly connected to a public water supply in urban areas of Nigeria. Moreover, there are many households in growing urban centers that usually rely on purchases from water merchants and water tanker owners. The source of that water cannot be guaranteed. It is collected from unprotected wells and streams (see Folasade Iyun and Adewale Oke, 2000).

Results for Nigeria show that the children who are born in a hospital are at lower risk of fever. However, the place of delivery is not associated with diarrhea or cough. For Egypt, the place of delivery has no effect on the diseases of children.

In spite of the fact that the data for both countries indicate that the vaccination status has mostly significant effect on the morbidity of children, the effect is positive! A reason for that could be because the high percentage of missing values associated with this variable could affect the results, or maybe the children have been vaccinated against other kind of diseases instead of the three types of diseases included in this application.

Metrical Covariates

In Egypt and Nigeria, childhood morbidity is associated with the child's age, mother's BMI and the mother's age at birth. The effect of mother's age in both countries is comparably higher in the young mothers ($< 20 - 22$). In other words, children from younger mothers are at higher risk, compared to who are from mothers in middle age (20-35 years).

The effects of child's age indicate a continuous worsening of diarrhea, fever and cough disease during the first 10-11 months of age and maybe even during the first 24 months of age in both countries. One reason for these results in both countries could be that there are some parents, as suggested by the child health literature, who prevented the breastfeeding shortly after birth and give their children various liquids instead of the mother's milk, which could lead to the infections. Other reasons for this could be also that there are some communities facing many problems which result children's diseases. These problems include lack of sanitation, access to clean water, municipal water range, unimproved water supplies (e.g wells, rivers, ponds, canals and unprotected springs) and lastly the unimproved sanitation for facilities such as holes, bushes and other places where human waste is not contained to protect it from contaminating the environment.

With regard to the effect of BMI, it has a slight effect on diarrhea and fever in Egypt, however the morbidity appears to worsen around the BMI of 30 until 35, and stabilizes after that. As for cough in Egypt, the effect of BMI is comparably high for mothers with BMI greater than 30. For Nigeria, the effect of BMI is associated with high risk of diarrhea for mothers with BMI less than 22, while it has a slight effect on fever and cough morbidity.

Spatial Effects

The estimates of the presumed spatial correlated district level random effects in fact showed strong evidence of spatial dependence in both countries.

In Egypt (figures 3.3, 3.7 and 3.11), there appear to be negative influences on child morbidity in the some provinces in Nile Delta, Upper Egypt and Sinai. The reasons for this high rates of morbidity in these areas could be.

Firstly, the level of the rehabilitant of the existing system and services might be low at these areas. Furthermore, the supply of water is available to 67% of the resident compared to 86%-99% in urban governorates and in lower Egypt resident (Abu Ali, 2002). Secondly, most of the poor were found in Upper rural Egypt which lead to the highest rate of illness, where 5.5 million poor people, out of the 10.7 million, live in these regions and 1.4 million poor people live in the urban parts of Upper Egypt (Poverty Reduction in Egypt Diagnosis and Strategy, 2002). Moreover, the report indicates that about 17% of the Egyptian population were poor in the year 2000. Thirdly, because of the lower standard of living in these areas, which has a direct impact on the rate of illiteracy and also on the educational level of mothers, leads to more poverty in these areas and lower level of sanitation and rehabilitations.

For Nigeria, the results show that the southeastern part of countries is associated with a higher risk of having the three types of diseases. In addition, some districts in the central and northern part of country are associated with high risk of fever and cough morbidity. The reason for this high rate of disease among these regions could be, the distribution of the socio-economic factors for these districts. For example, in some regions with significant disease risk, the risks could be caused by the high percentage of households which have no access to flush toilet or even have a lower level of sanitation and have no access to clean water. Based on the 2003 NDHS (Demographic Health Survey of Nigeria, 2003), we found that 54% of households in southeast having access for traditional toilet, 14.3% using flush toilet, and 28% having no access for any type of toilet and using bushes/fields as a toilet facility. In the northeast, more than 74% of households use traditional toilets. Previous studies reported that the southeastern regions are affected by a high level of pollution because these parts of the country have petroleum, associated with incessant oil spillage. For this reason, the pollution on these areas affected the health of children through the water and pollution that makes access to drinkable water sanitation difficult (Adebayo, 2002). Some other previous studies such as the NICS 2003 reported that the most important reason for why children were not fully immunized is "vaccine not available" in all geopolitical zones, except the south south and the northeast, or "the place of immunization is too far".

3.6 A Reanalysis Excluding Some Factors

As already mentioned in the preceding section, the childhood diseases are associated with some socio-economic factors, bio-demographic and health factors in both countries. Potential factors we considered in this section are child's sex, mother's age, child's age in months, mother's body mass index, mother's age at birth, mother's educational attainment (recoded to two categories; "incomplete secondary education and complete secondary school and higher" (reference category), whether mother had treatment during pregnancy, whether mother had antenatal care, source of drinking water, type of toilet (reduced to two categories; has flush toilet or others (reference category), locality where respondent lives, availability of radio and electricity in the household, and mother's current working and the geographical effects. This reanalysis is considered in this section because of the following reasons:

The included covariates in this section are used in the analyses have obtained in chapter 5 with probit model instead logit model. Therefore, we need to explore the difference between these results and the results have obtained later with probit model. Furthermore, the probit model is provided in chapter 5 due to make a comparison between the results of separate analysis using probit model with the results of latent variable models which use usually a probit model for the binary response variable (see chapter 5).

Results in this application are quite consistent to the results shown in the previous analysis.

Results of Egypt show that child's sex and the antenatal visit have a positive significant effect on the three types of diseases, whilst mother's current

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.855^*	0.265	-2.393	-1.842	-1.351
male	0.106^*	0.034	0.033	0.106	0.175
urban	-0.111^*	0.045	-0.203	-0.112	-0.022
work	0.019	0.044	-0.071	0.017	0.110
trepr	0.111^*	0.052	0.011	0.109	0.214
anvis	0.153^*	0.041	0.069	0.153	0.235
radio	-0.080	0.046	-0.167	-0.080	0.0133
elect	0.025	0.164	-0.316	0.017	0.329
water	0.029	0.052	-0.072	0.028	0.126
educ	-0.053	0.042	-0.132	-0.053	0.027
toilet	-0.085	0.080	-0.238	-0.085	0.087

Table 3.20: Fixed effects on diarrhea- Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.632^*	0.201	-1.032	-0.631	-0.229
male	0.074^*	0.029	0.016	0.073	0.135
urban	0.009	0.035	-0.059	0.0091	0.079
work	0.083^*	0.038	0.009	0.082	0.156
trepr	0.0477	0.047	-0.052	0.0495	0.145
anvis	0.133^*	0.034	0.072	0.134	0.208
radio	-0.106^*	0.039	-0.181	-0.106	-0.027
elect	-0.150	0.131	-0.414	-0.151	0.112
water	0.060	0.045	-0.025	0.059	0.152
educ	-0.047	0.036	-0.115	-0.048	0.025
toilet	-0.072	0.074	-0.210	-0.076	0.085

Table 3.21: Fixed effects on fever- Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.732^*	0.223	-1.144	-0.736	-0.311
male	0.079^*	0.031	0.020	0.080	0.138
urban	0.071	0.038	-0.002	0.071	0.150
work	0.113^*	0.039	0.034	0.115	0.192
trepr	0.043	0.049	-0.054	0.044	0.144
anvis	0.117^*	0.037	0.047	0.117	0.190
radio	-0.095^*	0.040	-0.174	-0.096	-0.017
elect	-0.028	0.145	-0.297	-0.030	0.260
water	-0.034	0.0464	-0.124	-0.034	0.066
educ	-0.062	0.036	-0.135	-0.062	0.005
toilet	-0.094	0.072	-0.227	-0.096	0.054

Table 3.22: Fixed effects on cough- Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-2.42^*	0.325	-3.07	-2.441	-1.80
male	0.082	0.046	-0.011	0.082	0.177
urban	-0.072	0.062	-0.183	-0.071	0.061
work	0.026	0.05	-0.071	0.026	0.122
trepr	0.062	0.057	-0.056	0.062	0.176
anvis	-0.105	0.057	-0.217	-0.104	0.014
radio	-0.074	0.055	-0.189	-0.070	0.029
elect	0.051	0.060	-0.069	0.050	0.1652
water	-0.072	0.074	-0.221	-0.075	0.077
educ	-0.020	0.089	-0.196	-0.018	0.147
toilet	-0.080	0.102	-0.303	-0.083	0.120

Table 3.23: Fixed effects on diarrhea- Nigeria.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.035^*	0.216	-1.466	-1.028	-0.622
male	0.028	0.039	-0.04	0.027	0.107
urban	-0.080	0.054	-0.185	-0.082	0.029
work	0.039	0.043	-0.047	0.039	0.121
trepr	0.116^*	0.052	0.0178	0.119	0.215
anvis	0.029	0.050	-0.068	0.029	0.124
radio	-0.045	0.048	-0.138	-0.046	0.049
elect	-0.018	0.052	-0.117	-0.017	0.076
water	0.036	0.062	-0.082	0.036	0.158
educ	0.038	0.063	-0.082	0.038	0.150
toilet	-0.198^*	0.082	-0.357	-0.195	-0.044

Table 3.24: Fixed effects on fever- Nigeria.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.26^*	0.24	-1.79	-1.25	-0.80
male	0.008	0.042	-0.074	0.008	0.090
urban	-0.104^*	0.055	-0.209	-0.102	-0.005
work	0.024	0.048	-0.066	0.020	0.125
trepr	0.182^*	0.056	0.076	0.178	0.293
anvis	0.050	0.054	-0.070	0.052	0.150
radio	0.011	0.050	-0.091	0.013	0.105
elect	0.003	0.054	-0.103	0.003	0.112
water	0.011	0.062	-0.105	0.011	0.13
educ	-0.01	0.070	-0.147	-0.007	0.124
toilet	0.007	0.080	-0.144	0.006	0.159

Table 3.25: Fixed effects on cough- Nigeria.

working is associated positively with diseases of fever and cough, and availability of radio is associated negatively with both diseases. In addition, the children who living in urban area are less likely to have diarrhea disease, but children from mothers had treatment during pregnancy are more likely to get diarrhea (tables 3.20 through 3.22). Note that variables such as educational level become non-significant in this application and are seen only at borderline. For Nigeria, the results show that most of covariates have either nonsignificant or slight effect on diarrhea disease. The treatment during pregnancy has a positive significant effect on fever and cough. As for children living in urban area, they are less likely to have cough disease, and the children living at household having flush toilet are less likely to have fever disease in Nigeria. On the other hand, factors place of deliver, interval birth, birth order and the vaccination status are excluded in this section. That is because their slight effects on the childhood diseases in both countries as shown in the previous results (tables 3.23 through 3.25).

Comparison of figure 3.13 with figures 3.1, 3.5 and 3.10 shows that the nonlinear patterns of the three types of diseases risk are very similar after exclusion some of the socio-economic factors and recoding some other covariates. The same is true for the spatial effects (figure 3.15). It shows very similar aspects to the analysis in earlier sections.

The same conclusion is true for non-linear and spatial effects in Nigeria.

3.7 Summary and Concluding Remarks

This chapter presents analysis which are investigated the effect of socio-economic, bio-demographic and health factors on childhood diseases in Egypt and Nigeria. Determinants that explain the levels of disease in both countries have been explored using geoadditive logit models. The analysis shows that male, place of residence, antenatal visit during pregnancy, having treatment during pregnancy, mother's current working, availability of radio are the importance factors which affect the health of children in Egypt. The educational level of the mother could also be importance for the child's health in Egypt. For Nigeria, the place of residence, type of toilet, treatment during pregnancy, source of water (on cough), educational level (slight effect) and place of delivery have an effect on the health of children.

Concerning the non-linear effects, a major finding for both countries is that the health status worsens until 11 months of age. The effect of BMI is slight on fever and cough morbidity in Nigeria, but seem to have significant effect (mothers with BMI < 22) on diarrhea disease. In contrast, it has a slight effect on diarrhea and fever morbidity in Egypt. However, it seems comparably higher for mothers with BMI $>$ 30-35 on the diseases. The effect of younger mother's ($<$ 22 years) is considerably high in both countries compare to their counterparts (see section 3.4). We found convincing and sizeable spatial effects in the both countries. The spatial effects suggest that Upper Egypt and some rural provinces in Nile Delta are affected by the childhood diseases. The situation in Nigeria reflects a higher risk of disease in southeastern areas through north-eastern parts of the country.

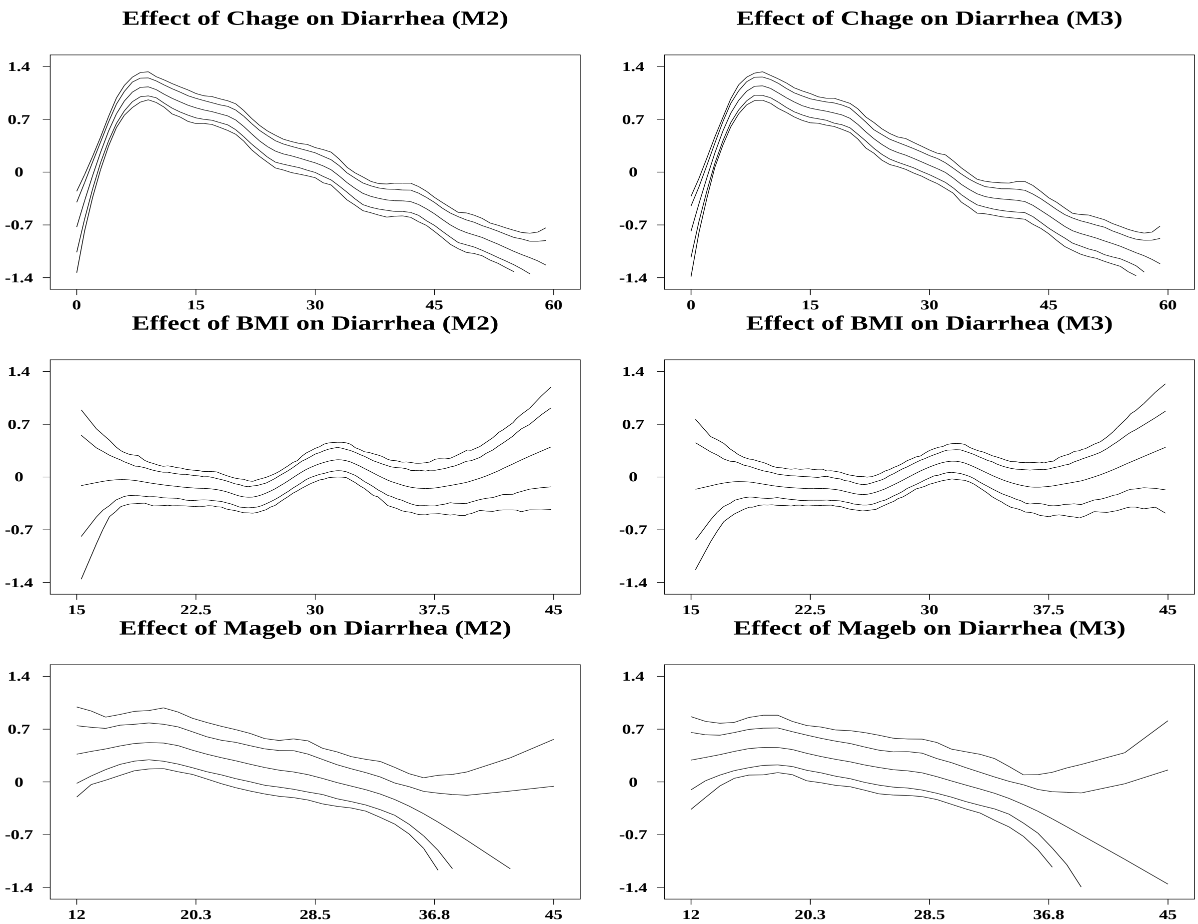


Figure 3.1: Non-linear effects from top to bottom: child's age, mother's BMI, and mother's age (for model M2-left panels), child's age, mother's BMI, and mother's age (for model M3-right panels) for diarrhea in Egypt.

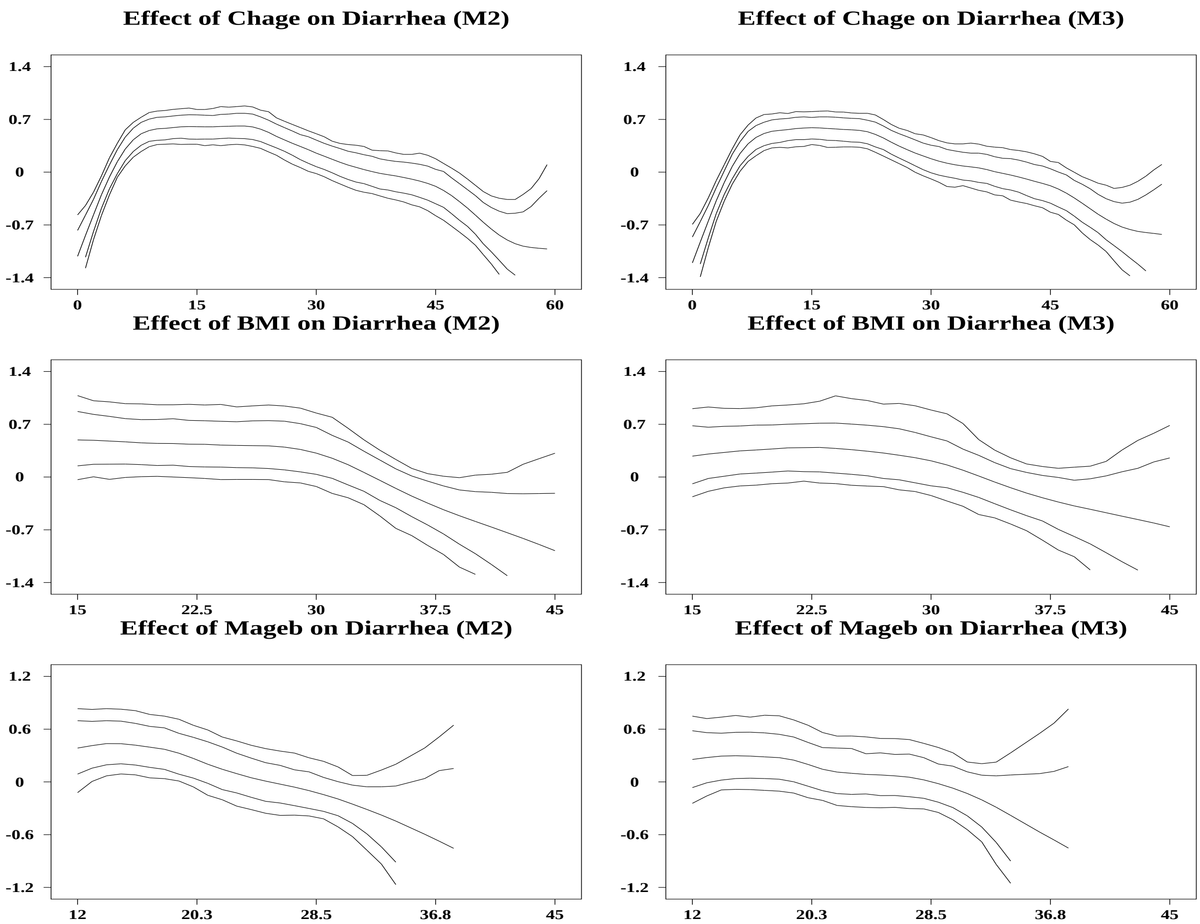


Figure 3.2: Non-linear effects from top to bottom: child's age, mother's BMI, and mother's age (for model M2-left panels), child's age, mother's BMI, and mother's age (for model M3-right panels) for diarrhea in Nigeria.

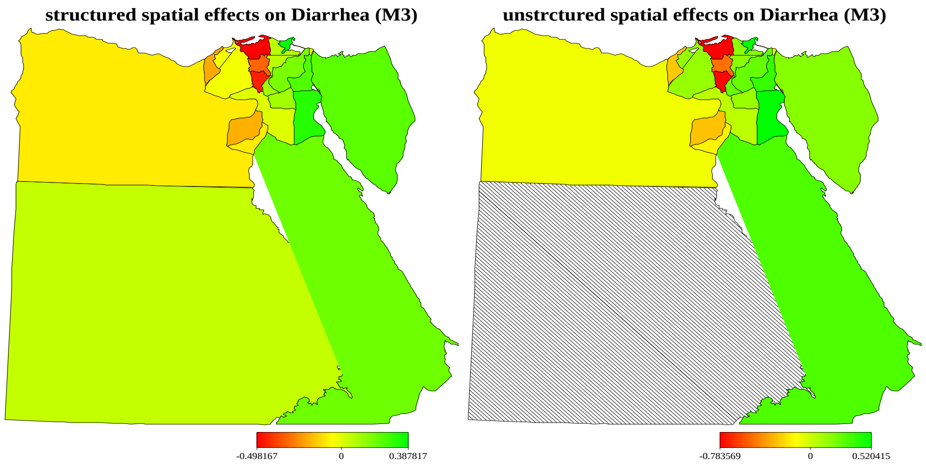


Figure 3.3: Maps of Egypt for diarrhea showing structured (top left) and unstructured (right left) spatial effects in model M3.

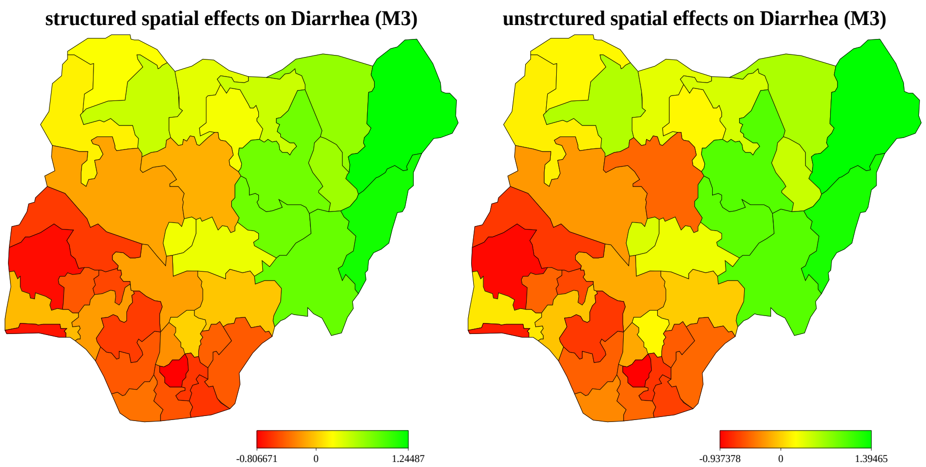


Figure 3.4: Maps of Nigeria for diarrhea showing structured (top left) and unstructured (right left) spatial effects in model M3.

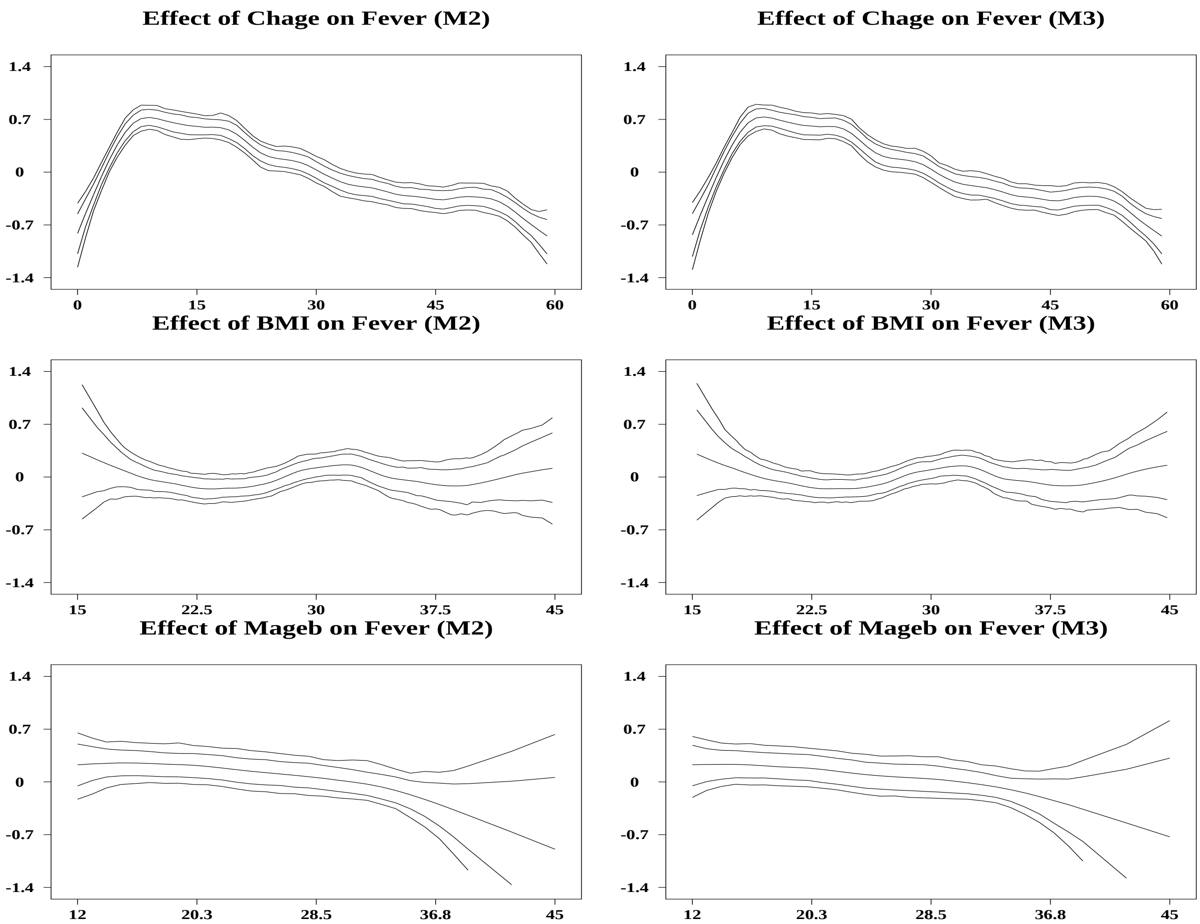


Figure 3.5: Non-linear effects from top to bottom: child's age, mother's BMI, and mother's age (for model M2-left panels), child's age, mother's BMI, and mother's age (for model M3-right panels) for fever in Egypt.

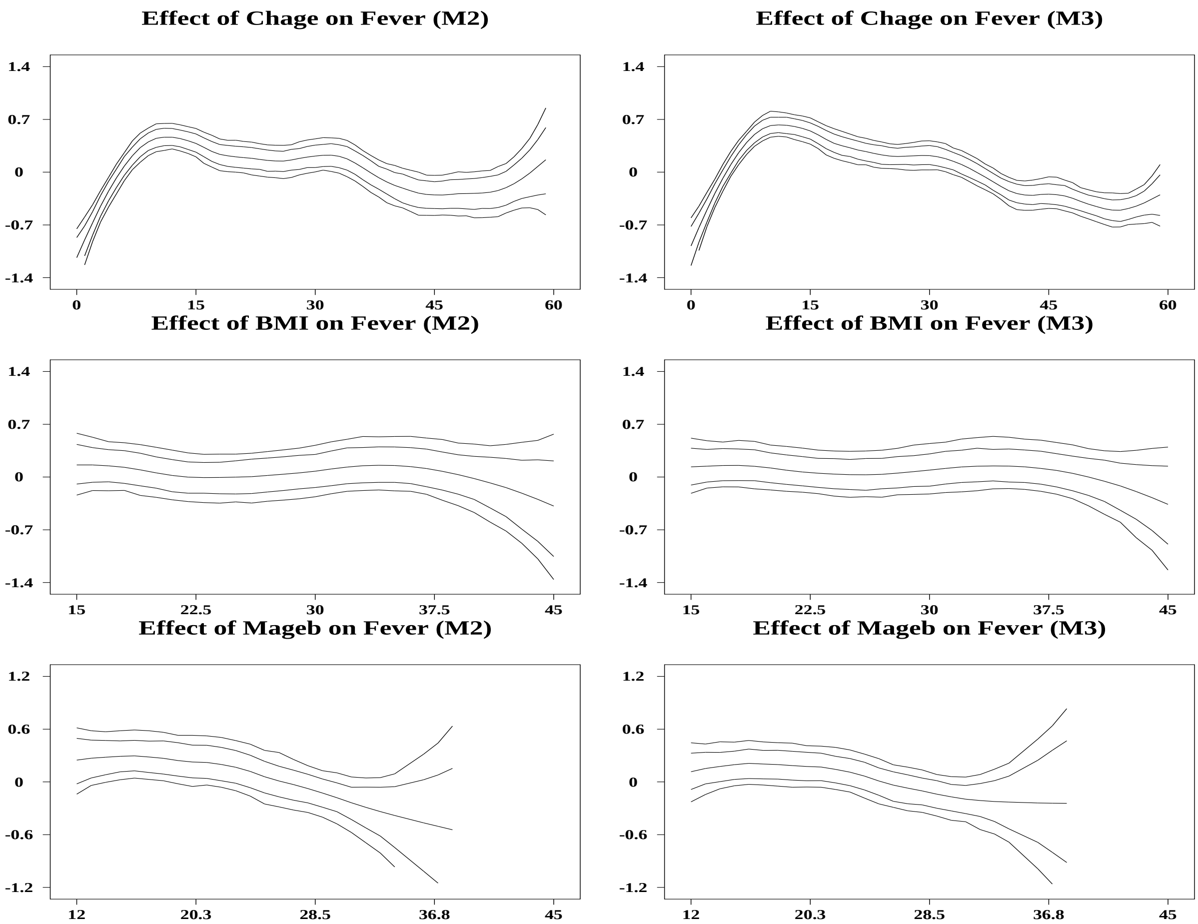


Figure 3.6: Non-linear effects from top to bottom: child's age, mother's BMI, and mother's age (for model M2-left panels), child's age, mother's BMI, and mother's age (for model M3-right panels) for fever in Nigeria.

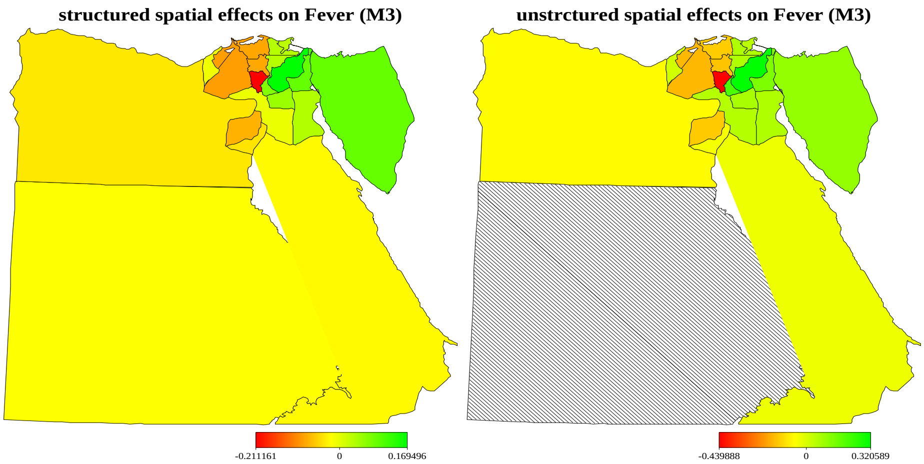


Figure 3.7: Maps of Egypt for fever showing structured (top left) and unstructured (right left) spatial effects in model M3.

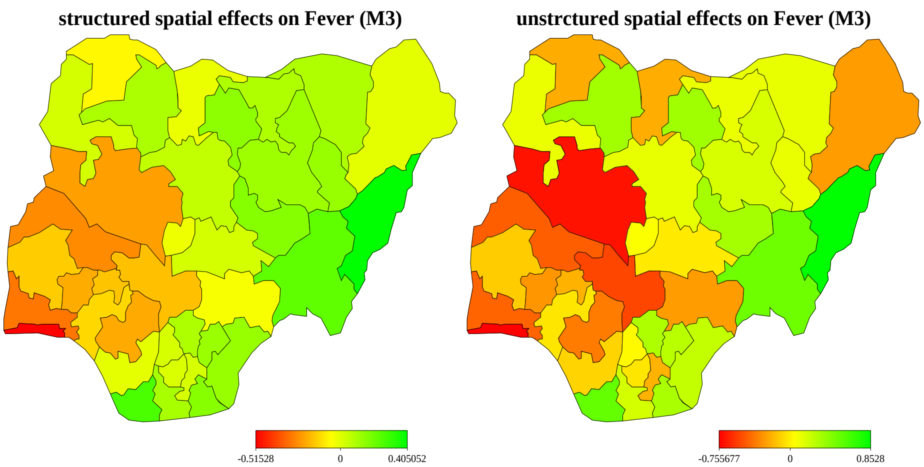


Figure 3.8: Maps of Nigeria for fever showing structured (top left) and unstructured (right left) spatial effects in model M3.

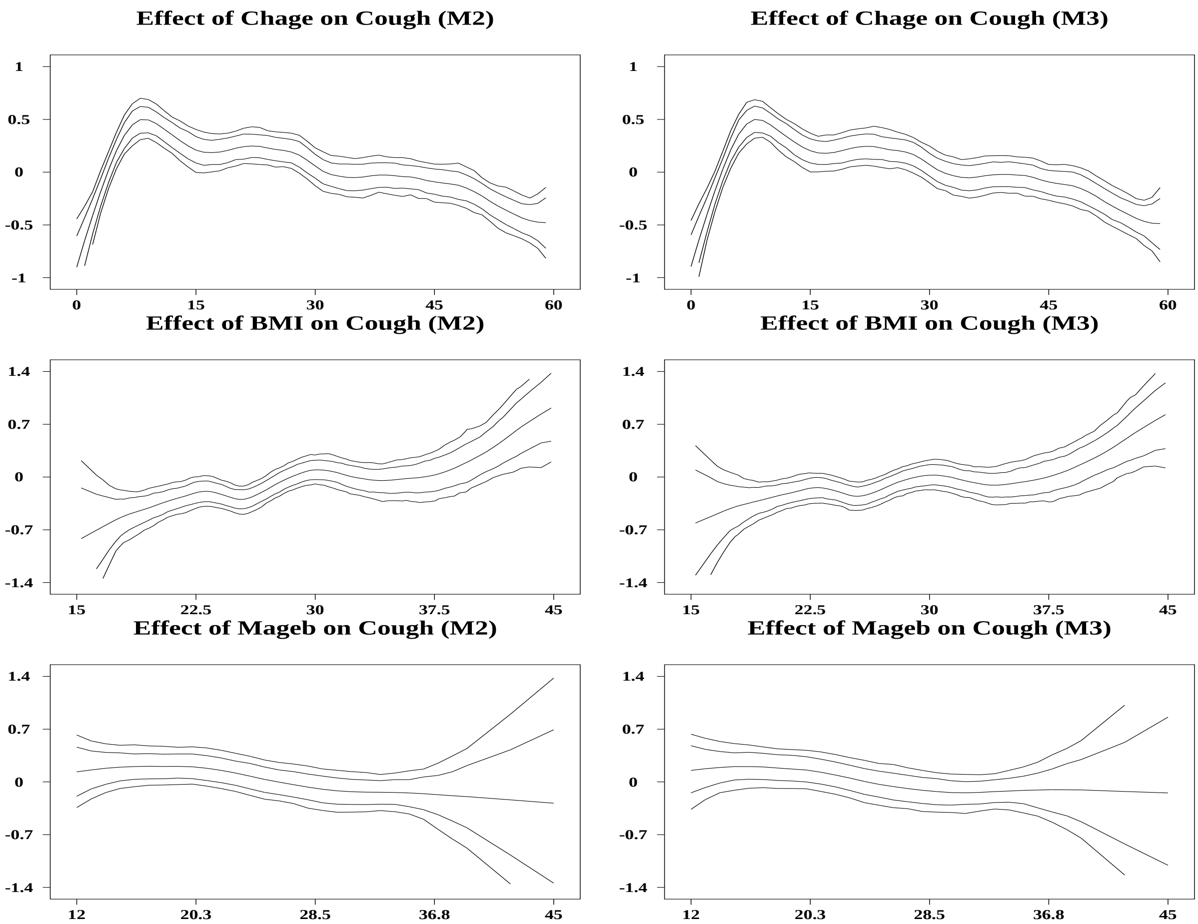


Figure 3.9: Non-linear effects from top to bottom: child's age, mother's BMI and mother's age (for model M2-left panels), child's age, mother's BMI, and mother's age (for model M3-right panels) for cough in Egypt.

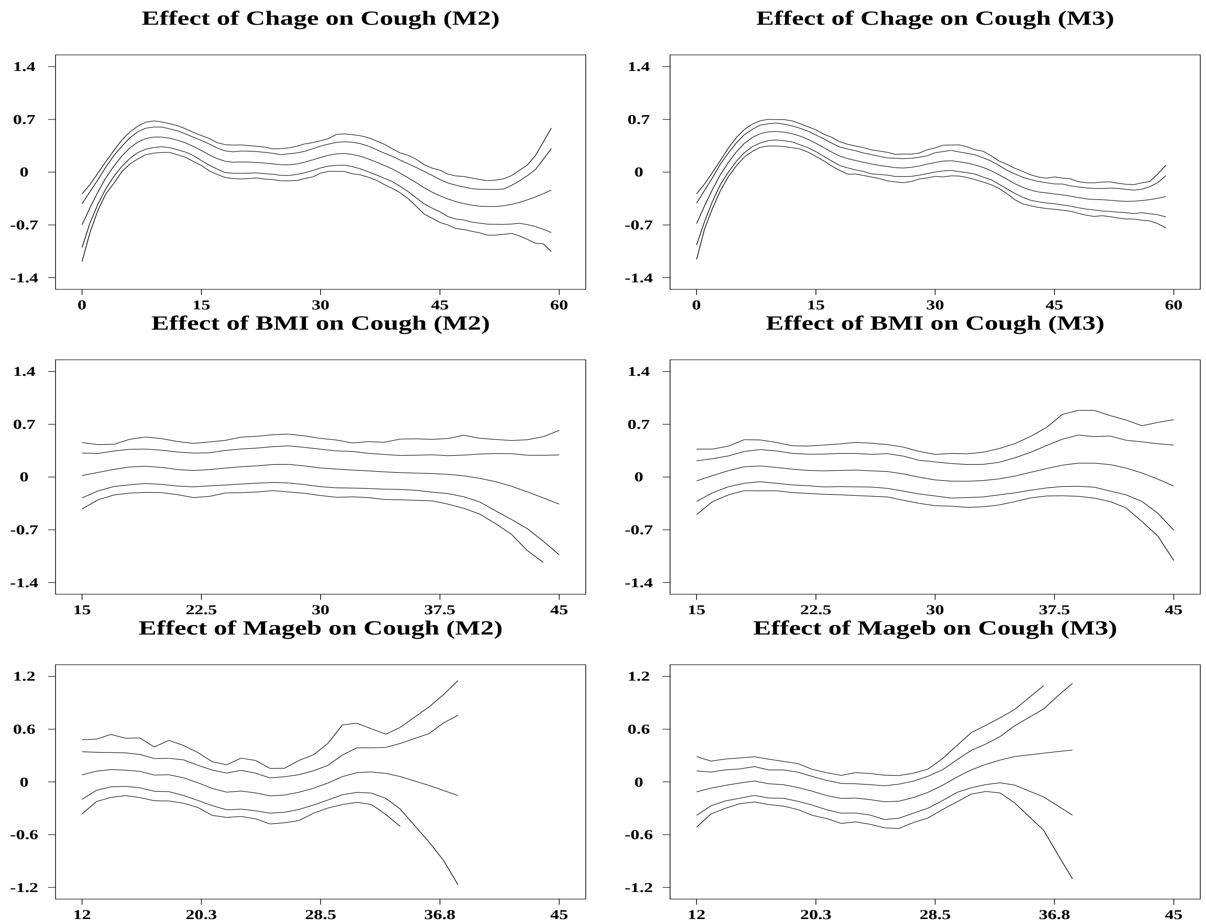


Figure 3.10: Non-linear effects from top to bottom: child's age and mother's BMI, and mother's age (for model M2-left panels), child's age and mother's BMI, and mother's age (for model M3-right panels) for cough in Nigeria.

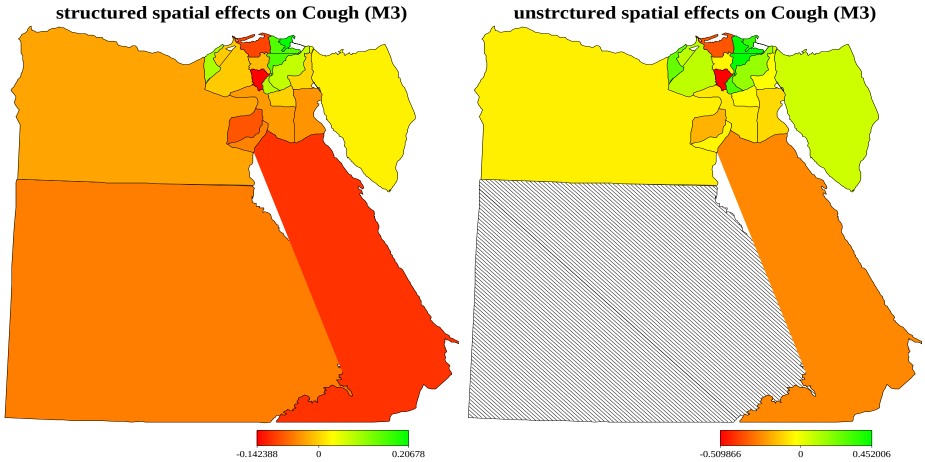


Figure 3.11: Maps of Egypt for fever showing structured (top left) and unstructured (right left) spatial effects in model M3.

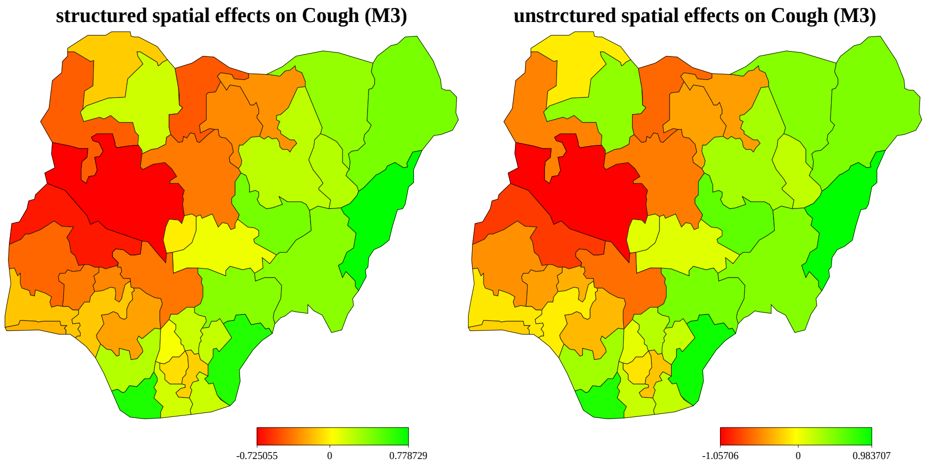


Figure 3.12: Maps of Nigeria for cough showing structured (top left) and unstructured (right left) spatial effects in model M3.

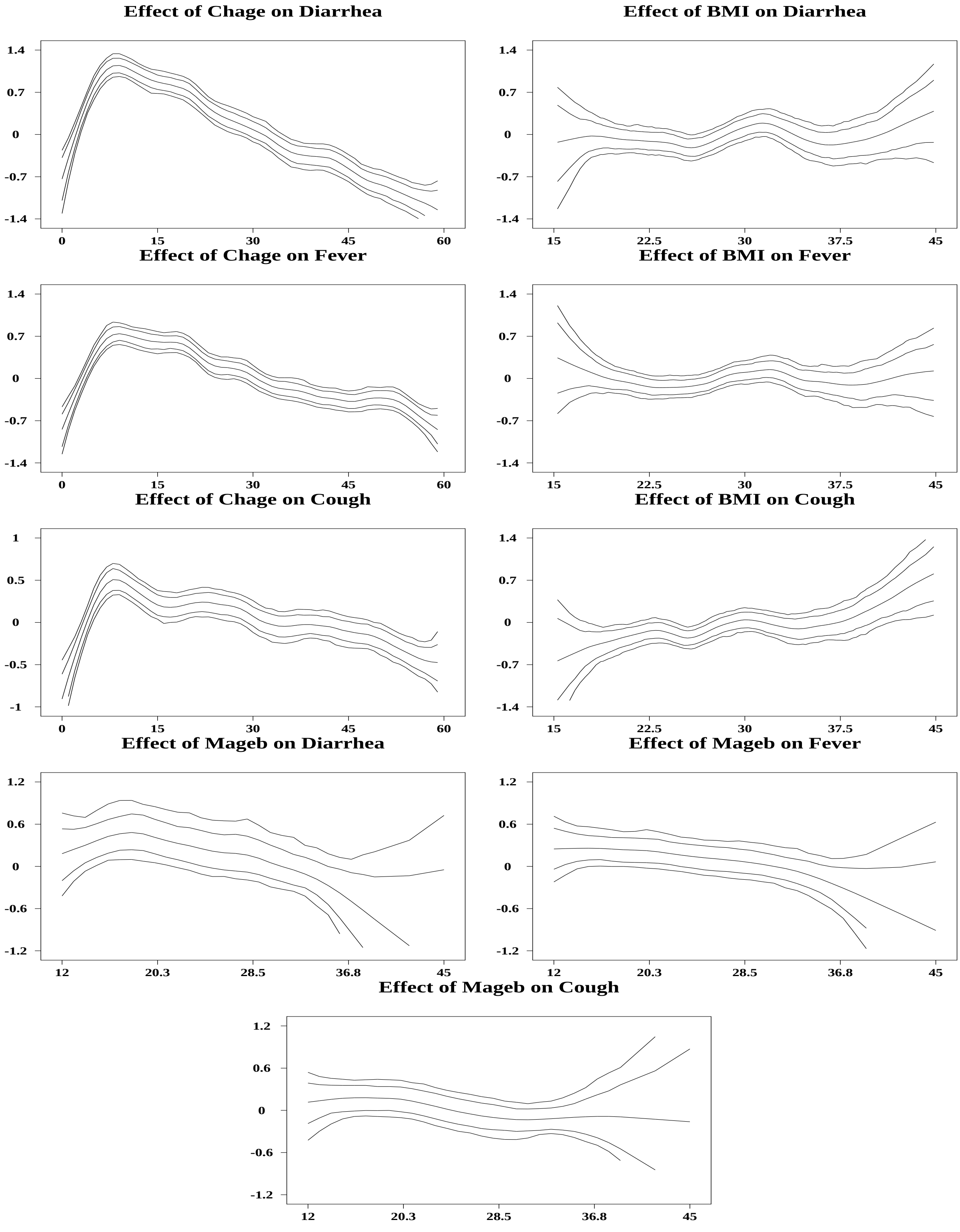


Figure 3.13: Non-linear effects from top to bottom: child's age, mother's BMI and mother's age for diarrhea, fever and cough diseases, respectively in Egypt.

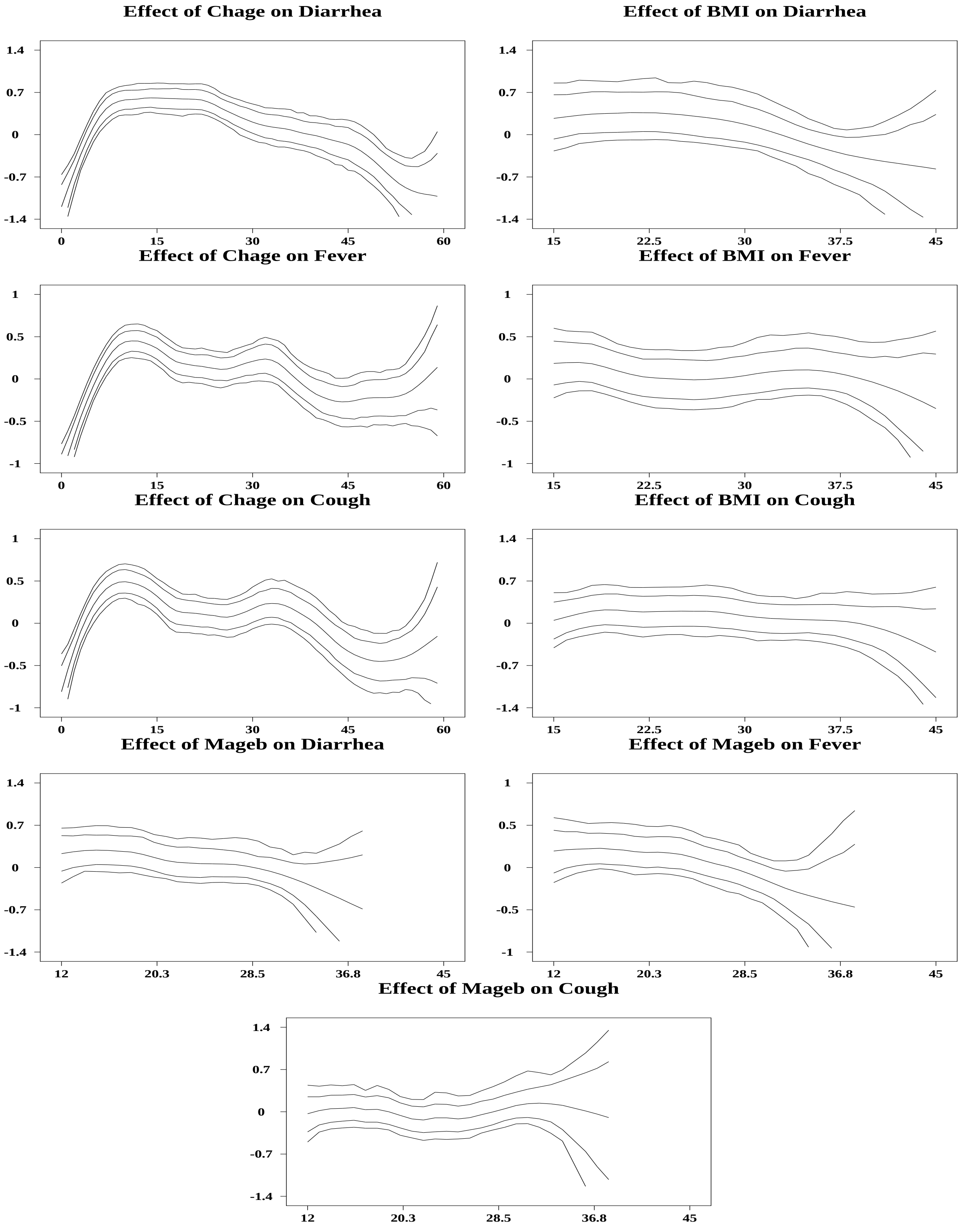


Figure 3.14: Non-linear effects from top to bottom: child's age, mother's BMI and mother's age for diarrhea, fever and cough diseases, respectively in Nigeria.

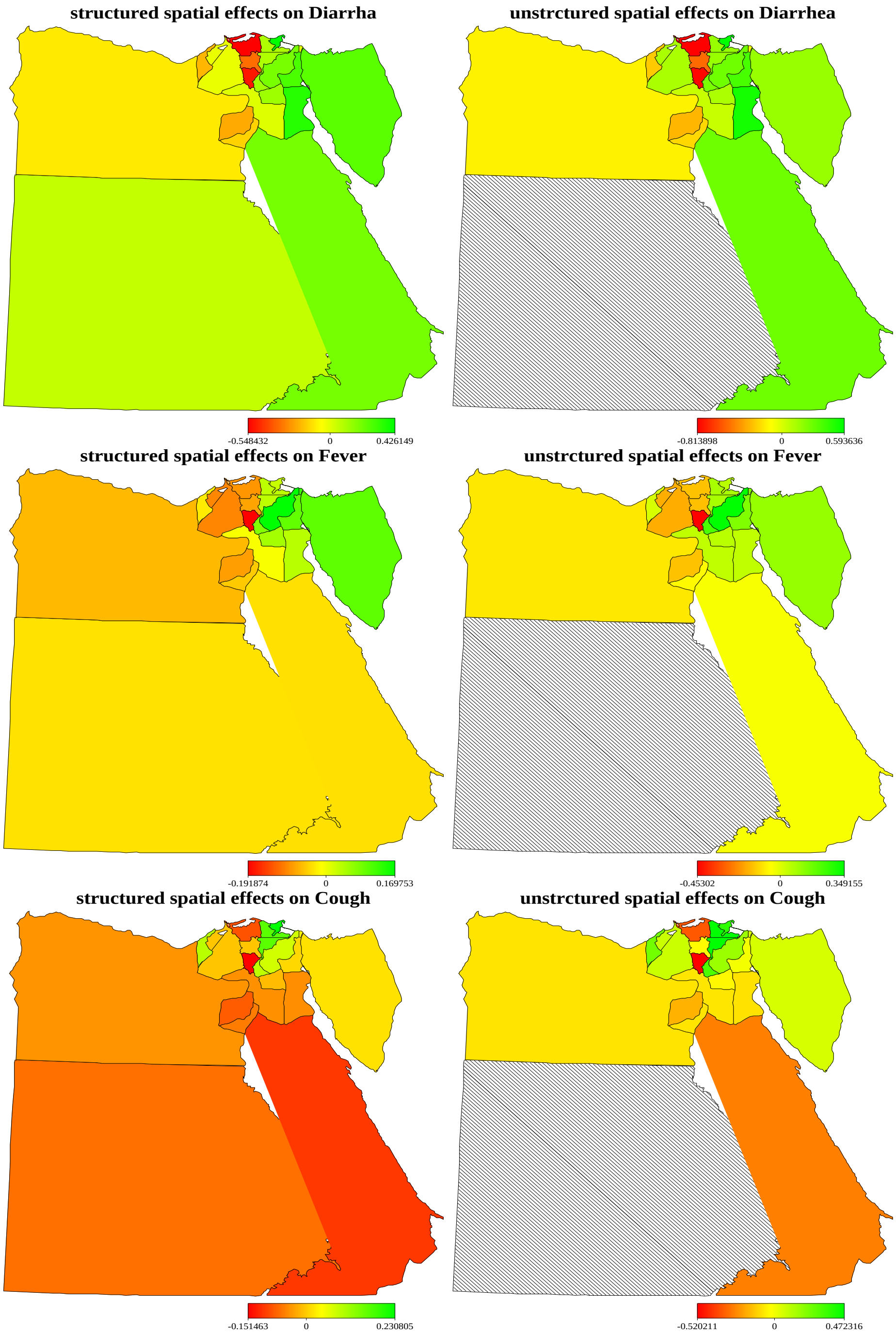


Figure 3.15: Maps of Egypt for diarrhea, fever and cough diseases (from top to bottom) showing structured (left panels) and unstructured (right panels).

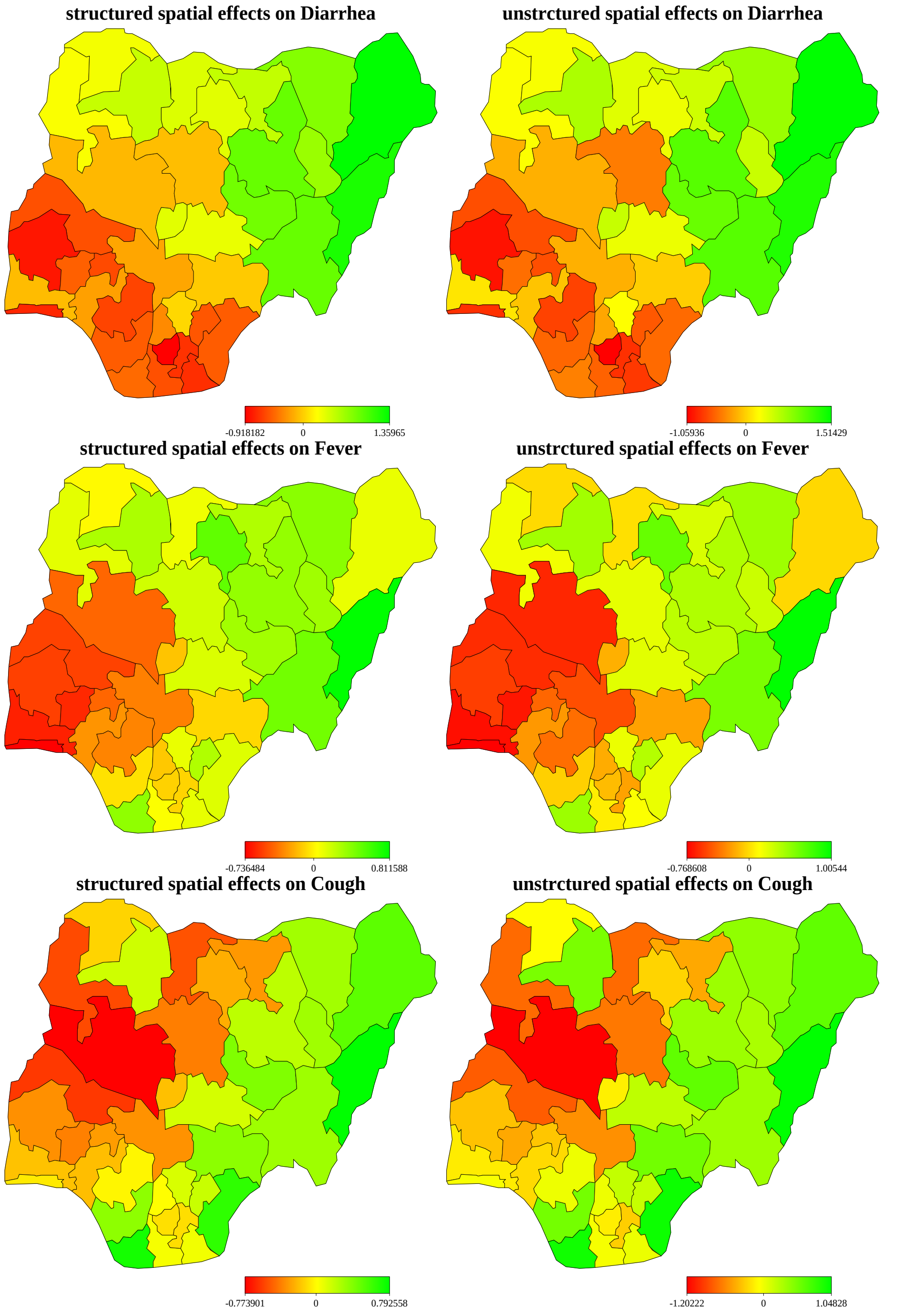


Figure 3.16: Maps of Nigeria for diarrhea, fever and cough diseases (from top to bottom) showing structured (left panels) and unstructured (right panels).

Chapter 4

Latent Variable Models

Abstract

Latent variable models (LVM) are used successfully to explain the interrelationships between the components of multivariate observable responses, to measure underlying unobservable constructs, and to assess the influence of covariates on observable and latent variables. In this chapter we introduce the basic idea behind the latent variable models (LVM) for various types of response variables such as binary, contentious or ordinal responses, where covariate effects on the continuous latent variables are modelled using a flexible geoadditive predictor.

4.1 Basic Ideas of Latent Variable Models

Latent variable models provide an important tool for the analysis of multivariate data. When the joint distribution of a set of random variables is specified by a statistical model it becomes a latent variable model if some of them are unobservable.

There are many reasons for why latent variables might be introduced into a model in the first place and how their presence contributes to statistical investigation. One reason is to reduce dimensionality. That means, if the information contained in the interrelationships of some variables can be useful, to an approximation, in a much smaller set, that will improve the ability to see structure in the data. That is the idea which is behind a lot of factor analysis models and more recent applications of parametric structural models. Secondly, latent variable models play a prominent role in many fields to which statistical methods are applied. Some of these fields are social science, psychology and politics. There are two sorts of variables to be considered in terms of latent variable models: variables which can be directly observed, known as manifest variables, and latent variables, which cannot be measured directly.

Many constructs that are of interest to social scientists cannot be observed directly. Examples are preferences, attitudes, behavioural intentions, and personality traits. Such constructs can solely be measured indirectly by means of observable indicators, such as questionnaire items designed to elicit responses related to an attitude or preference. There are various types of scaling techniques which have been developed for deriving information on unobservable constructs of interest from the indicators. A latent variable model can be a nonlinear, path analysis or graphical model. In addition the manifest variables, the model can include one or more unobserved or latent variables which represent the constructs of interest. There are two various assumptions defining the causal mechanisms underlying the responses. The first one assumes that the responses on the indicators are the result of an individual's position on the latent variable. The second one is that the manifest variables have nothing in common after controlling for the latent variable. This is usually referred to as the principle of local independence. The two remaining assumptions concern the distributions of the latent and manifest variables. Depending on these assumptions, one obtains different kinds of latent variable models. According to Bartholomew (1995), the four main kinds are as follows:

factor analysis (FA), latent trait analysis (LTA), latent profile analysis (LPA), and latent class analysis (LCA) (see figure 4.1).

In factor analysis (FA) and latent trait analysis (LTA), the latent variables are treated as continuous normally distributed variables. However, in latent

manifest variable / latent variable	Metrical	Categorical
Metrical	Factor analysis	Latent trait analysis
Categorical	Latent profile analysis	Latent analysis

Figure 4.1: Classification of LVM (adapted from Bartholomew, 1995)

profile analysis (LPA) and latent class analysis (LCA), the latent variable is discrete, and assumed to come from a multinomial distribution.

In most cases, their conditional distribution (given the latent variables) is assumed to be normal. In 4.1 LTA and LCA, the indicators are dichotomous, ordinal, or nominal categorical variables, and their conditional distributions are assumed to be binomial or multinomial. The more fundamental distinction in Bartholomew's typology is the one between continuous and discrete latent variables. A researcher has to decide whether it is more natural to treat the underlying latent variable as continuous or discrete. However, as shown by Heinen (1996), the distribution of a continuous latent variable model can be approximated by a discrete distribution. This shows that the distinction between continuous and discrete latent variables is less fundamental than one might initially think. The distinction between models for continuous and discrete indicators turns out not to be fundamental at all. The specification of the conditional distributions of the indicators follows naturally from their scale types. The most recent development in latent variable modelling is to allow for a different distributional form for each indicator. These can, for example, be normal, log-normal, gamma, or exponential distributions for continuous variables, binomial for dichotomous variables, multinomial for ordinal and nominal variables, and poisson, binomial, or negative-binomial for counts. Depending on whether the latent variable is treated as continuous or discrete, one obtains a generalized form of LTA or LCA.

The main purpose of factor analysis is to determine the correlations between a set of observed variables that can be interpreted by a few numbers of latent variables, and how that could be identified. The factor analysis model can be found in two ways

1- The model which allows for ordinal or binary indicators. Typically researches have used ordinal data in classic factor analysis models, which were assumed to be normally distributed.

2- A latent variable model including covariates which influence the indicators or the latent variables. Most statistical studies assume that the influence of the covariates on the indicators and the latent variable as a strictly linear influential.

The original form of factor analysis, was inverted in the Psychology field by the British Psychologist Spearman in 1904. He hypothesized that performance for each set of intellectual tasks sharing with performance and all other intellectual tasks, the general intellectual ability cannot be directly obtained, and therefore there is a need for a latent variable.

The LVM presented in this work includes binary and continuous indicators.

In this work, the LVM depends on a Bayesian approach where all unknown population parameters are considered as random.

In order to understand the idea of LVM, we have to distinguish between two types of variables: The observable variables which are called indicators or manifest variables, and the unobservable variable which is called latent variable.

LVMs are mostly used in the fields of psychology and stoical science. That is because the interesting variables in these areas cannot be directly measured using the traditional statistical methods, and thus are represented by latent variables. These models (LVMs) are also used in the field of medicine, where patients suffer from disease of syndromes which is a variety of effects such as Fetal Alcohol syndrome, and Downs Syndrome, which are assumed to be indicators in many teratology studies (Holmes et al., 1987).

4.1.1 Notation and General Formulation

Let y' be a vector of p manifest variables (or indicators) which are denoted by: $y_0 = (y_1, y_2, ..., y_p)$. The latent variables are denoted by a $q \times 1$ vector υ =$(\upsilon_1, \upsilon_2, ..., \upsilon_q)$ where $q < p$. One wants to find a set of latent factors $(\upsilon_1, \upsilon_2, ..., \upsilon_q)$ with a smaller number of components $q < p$ than the observed variables that contain essentially the same information. If both the response variables and the latent factors are normally distributed with zero means and unit variance and without covariates, this leads to classic factor model (see Jöreskog, 1979). We will distinguish between two different sets of covariates.

- covariates that affect the indicators directly $w' = (w_1, w_2, ..., w_k)$
- covariates that affect the indicators indirectly $x' = (x_1, x_2, ..., x_r)$

Covariates can be of any type, such as metrical, categorical (dummy variables) or ordinal.

4.1.2 Latent Variable Models (one factor) without Covariate Effects

Here, we briefly discuss the types of models that will be studied in this work. figure 4.2 shows the relationships that are allowed to be modelled using three response variables, which could be binary or continuous and one factor. It shows that the three observed variables $y' = y_1, y_2, y_3$ are indicators of a single latent variable υ_1. In addition, the individual error terms added to the observed manifest variables and illustrated by arrows without origins. Furthermore, the basic idea of latent variable models or the factor analysis

(Spearman, 1904) is that the multidimensional vector of p manifest variables y can be represented by one or more latent factor υ with a lower dimension of q. Consequently factor analysis reduces the dimensionality of the data in such a way that the interrelationships among the observed variables are preserved as much as possible.

The basic factor analytic model for Gaussian response consists of so-called *measurement model*

$$y_i = \Lambda \upsilon_i + \varepsilon_i, \tag{4.1}$$

$$\upsilon_i \sim N(0, I), \;\; \varepsilon_i \sim N_p(0, \Sigma),$$

for each observation i. The factor loadings Λ is a $p \times q$ dimensional matrix of regression coefficient which is called factor loadings that indicate the relationship between the latent variable υ, and the indicators (manifest variable) y_i. The term ϵ_i represents a p-dimensional error term.

4.1.3 Linear latent Variable Models (one factor) with Covariate Effects

The reasons why we need to extend the basic factor model are as follows; On one hand, it is useful to include the explanatory variables (named direct effects w) which affect the observed variables directly. On the other hand, it is interesting to know how the explanatory variables modify the latent factor, and hence affect the observed variables indirectly (those are called indirect effects x). This work focuses on both types of exploratory variables (direct and indirect effects) as we are interested in how variables, e.g demographic variables, affect the latent variables. In the recent applications, some of those variables might not exist. For example, there could be a case where only exploratory variable affect the latent variables or exploratory variable that only affect the indicators (manifest variables). Most structural modelers following Jöreskog, distinguish between two conceptually distinct parts of latent models, namely a structural part and a measurement part. The structural part of a model specifies the relationships among the latent variables and the measurement part specifies the relationship of the latent to the observed variables. In other words, a linear latent variable model consists of two parts:

Firstly, the part that accommodates the effect of the latent variables and a set of observed covariates on the indicators. It is called the *measurement model* (with direct effects).

$$y_i = \Lambda \upsilon_i + A w_i + \varepsilon_i. \tag{4.2}$$

These w_i are direct effects which directly affect the observed manifest variables and A is the matrix of regression coefficients. Secondly, the part of the model that links a set of observed covariates with the latent variables, called the *linear structural model.*

$$\upsilon_i = \gamma x_i + \zeta_i. \tag{4.3}$$

These x_i are indirect effects which modify the latent factors, and hence affect the observed variables. The matrix γ contains the regression coefficients of the indirect covariates x. Figure 4.3 shows that the latent variable υ_1 and the observed variable w_1, accounting for the associations among the y variables. The direct arrow from w_1 to y_1 allows the mean level (thresholds) for variable y_i to be different for different values of the w_1 variable. Finally, $x' = x_1,\ x_2$ have an effect on the latent variable υ_1. Note that variable x needs to be different from variable w for identification reasons.

4.1.4 Underlying Variable and Item Response Theory

There are two main approaches for conducting latent variable analysis often used in the applications.

- One is the underlying variable approach (UVM) developed within the structural equation modelling framework (SEM). It is supported by LISREL software (Jöreskog and Sörbom, 19999), EQS (Bentler, 1992),

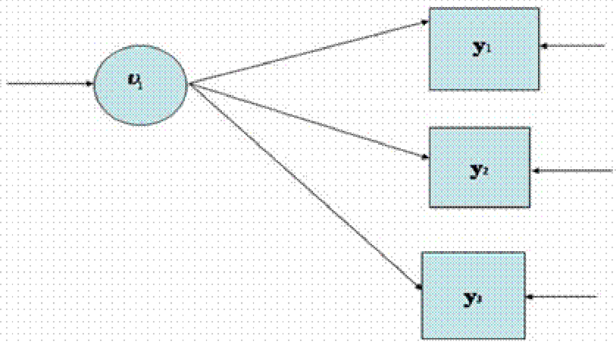


Figure 4.2: Path diagram of the latent variable models without covariate effects

and Mplus (Muthen, 2000). It is assumed by the paradigm that the categorical observed variable y_i are created by an underlying unobserved continuous variable which is distributed by normal distribution. The model in this work uses the UVA for the factor analysis of binary and continuous indicators. That is because the underlying normally distributed variables can be quite naturally incorporated into a Bayesian estimation approach.

- The other approach is the Item Response Theory approach (IRT). The Item Response Theory approach specifies the conditional distribution of the complete p-dimensional response variable as a function of the

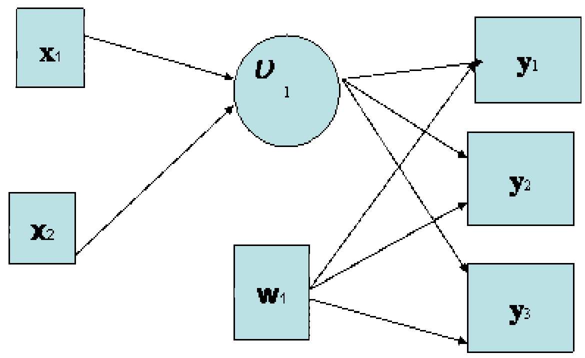


Figure 4.3: Path diagram of the latent variable model including covariate effects (structural equation).

latent variable/ explanatory variables. In the Item Response Theory approach, the element of analysis is a whole response pattern of the sample members. These assumptions of IRT are made that of the conditional independency (responses to the p ordinal) indicators are independent conditionals on the latent variable υ, the set of explanatory variables x, and the multinomial assumption for the conditional distribution. Furthermore, the latent variables are assumed to be normally distributed. Within the IRT framework, the correlated latent variables can be also fitted (see Jöreskog and Moustaki, 2001).

The log likelihood and maximum likelihood with an EM algorithm are used in the IRT framework. Verhelst, Glas, and Verstralen (1994), Zwinderman (1997) and Glas (2001) have discussed the one parameter logistic model with covariates effects, Sammel, Ryon, and Legler (1997) have discussed a uni-dimensional latent trait model for binary and normal outcomes which allowing for covariate effects.

Furthermore, Moustaki (2003) has discussed a multi dimensional model for ordinal indicators with covariates effects.

A comparison between UVM and IRT models for ordinal indicators without covariates effects is reported by Moustaki (2000) and Jöreskog and Moustaki (2001).

All these papers which have been mentioned above, only consider the parametric effects in modifying the indicators and the latent variables.

However, we resolve these restriction in this work and introduce non-parametric effects on the latent variables see Raach (2005).

The non-parametric predictor that influenced the indicators directly might also include, a more detailed among the analyses of covariates on latent variables are considered as much more illuminating from an applied researcher's view.

4.1.5 Bayesian Approach to LVM

The traditional way to influence the latent variable is via the likelihood function and standard methods. There is very little work which has been done on Bayesian methods, because of the large number of parameters in typical models.

The Bayesian framework, has been rarely employed in applications until the early 1990s due to a lack of computing power and the suitable numerical methods (e.g. MCMC).

Here, we give an overview of Bayesian developments which have occurred in relevance to factor analysis using continuous and binary response variables depending on the LVM.

Bayesian models, are useful especially in using for problems which cannot be dealt with easily by other approaches, such as the estimation of factor analysis for multilevel binary responses (Ansori and Jedidi, 2000), dynamic factor components with time series (Aguilar and West, 2000) and determining the right number of latent variables (Lopes and West, 2004).

There are some reasons or advantages which make a Bayesian approach model useful in using in this thesis.

- The values of the latent variables (Factor scores) can be automatically estimated in the Bayesian approach framework. While these values have to be calculated separately after model estimation in other approachs.
- The marginal distribution of the parameters and the values of latent variables are obtained in a Bayesian approach framework. Hence, by looking at the marginal distributions, the uncertainty and the range of parameter values can be easily analyzed after the estimation process.
- The full posterior distribution of the model is analyzed, and hence the complete information among the indicators (manifest variables) is incorporated in the estimation process.

4.2 A Bayesian Geoadditive LVM

The LVM with covariates consists of two main approaches: the measurement model for continuous and binary response with covariates influencing the indicators directly (direct effects); and the structural model explaining the modification of the latent variables by covariates (indirect effects) (see Fahrmeir and Raach, 2006).

4.2.1 Measurement model

Underlying Variable Approach (UVA)

The binary variables y_{ij} are taken to be manifestations of some underlying continuous unobserved variables y^*_{ij}.

Each manifest variable $1 \leq j \leq p$ can be of continuous, binary (or ordinal) type. Where $1 \leq i \leq n$. The connection between the binary variable y_{ij} and the underlying variable y^*_{ij} is

$$y_{ij} = 1 \Longleftrightarrow y_{ij}^* > t_j.$$
$$y_{ij} = 0 \Longleftrightarrow y_{ij}^* \leq t_j$$

Because of the identification restriction, the t_j of all indicators j are fixed to zero and $var(\epsilon_i)$=1.

The essence of UV is to treat the y_i as generated by the classical factor analysis model.

The relationship between the y_i^* variables and the latent variables υ in the term of *measurement model* excluding direct effects is given by

$$y_{ij}^* = \lambda_0 + \Lambda \upsilon_i + \epsilon_{ij}, \quad \epsilon_{ij} \sim N_p(0, I).$$

where Λ is $p \times q$ matrix which is composed of the factor loadings, indicating strength of relationship between latent factors and indicators.

For continuous (Gaussian) indicators there is no need for underlying variable, so that

$$y_{ij}^* = y_{ij} \quad \epsilon_{ij} \sim N(0, \sigma_j^2),$$

The logistic distribution function could also be used instead of the standard normal distribution function; however we use the standard normal distribution function because the parameter estimates for both function lead to similar results in prediction (Moustaki, 2003).

Secondly, the relationship between the y_i^* variables and the latent variables υ_i in the term of *measurement model including direct effects* is given by

$$y_{ij}^* = \lambda_0 + \Lambda \upsilon_i + A w_i + \epsilon_{ij} \quad \epsilon_{ij} \sim N_p(0, \Sigma) \tag{4.4}$$

The direct covariates are summarized in the d-dimensional vector $w_i = (w_{i1}, .., w_{id})'$ and the $p \times q$-dimensional matrix A.

The direct effects provide additional information about data structure and increase the strength of dimensionality through the relationship between y^*_{ij} and w_i, used in the analyses later. Here ϵ_i is distributed normally $\epsilon_i \sim N_p(0, \Sigma)$ and $\Sigma = diag(\sigma_1^2, .., \sigma_p^2)$, υ is a $(1 \times p)$ latent variables that explain the relationships among the indicators. The $p \times q$ matrix Λ is the matrix of loading factors which indicate the relationship between the latent variables and the indicators, and λ_0 is the intercept.

In such models, the correlations between the y_i variables are explained by both latent variables and covariates, instead of the latent variable alone.

4.2.2 Structural Model

Here, the indirect effects are included to modify the latent variables by introducing the structural equation part of the model, i.e.

$$\upsilon_i = \eta_i^{geo} + \xi_i \tag{4.5}$$

where $\xi_i \sim N_q(0, I_q)$ and a geoadditive predictor $\eta_i^{geo} = (\eta_{i1}, \eta_{i2}, ..., \eta_{iq})$.

$$\eta_{ir}^{geo} = f_{r1}(x_{i1}) + ... + f_{rg}(x_{ig}) + f_{r,spat(s_i)} + \gamma_r' u_i \tag{4.6}$$

where g denotes the number of different nonparametric functions f_{rh} of metrical covariates $x_{ih}(1 \leq h \leq g)$, $f_{r,spat}$ is the spatial effect of the region s_i and γ_r is a vector of values in the $r - th$ row of the $q \times m$ matrix of γ of standard regression coefficients. As for u_i, it is a $m \times 1$ vector of fixed covariates of observation. If (4.6) does not contain the term of spatial effect f_{spat} and the covariates $x_{x_{ih}}$ are metrical, then the additive LVM is obtained. As in our case, where (4.6) contains a spatial effect f_{spat}, then a geoadditive LVM is obtained.

4.2.3 Identification Problems

There are two sources of identification problems.

First one is associated with modelling of ordinal variables, but our focus is on binary indicators in this thesis. Second is related to the uniqueness of factor loadings matrix Λ and factor scores.

For the binary indicators the t_j of all indicators j are fixed to zero and $var(\epsilon_i)$=1 in order to solve the identification problem. For more details see Raach, 2005.

Uniqueness of factor analysis and scores

$$y_i^* = \lambda_0 + \Lambda T^{-1} T \upsilon_i + A w_i + \varepsilon_i \tag{4.7}$$

Consider the transformation equation (4.4) with a $q \times q$ non-singular matrix T (e.g. Bartholomew, 1987), i.e.

where ΛT^{-1} is a loading matrix, new latent scores $T\ \upsilon_i$ and $V(\upsilon_i) = T\Psi T'$.

Without any restrictions for Λ or Ψ, a different number of models may be created. Since the matrix T consists of q^2 elements, then we have to set q^2 restrictions in the model. For this reason the latent scores have a standard normal distribution, and no correlations among the latent variables exist.

In the traditional exploratory factor analysis, the variance matrix of the latent scores can be chosen to be q-dimensional identity matrix I_q, leading to $\upsilon_i \sim N_q(0, I_q)$.

For this reason, the latent scores have a standard normal distribution, and no correlations among the latent variables could exist. The model is invariant under transformations with orthogonal $q \times q$ matrix V of form $\widetilde{\Lambda} = \Lambda V'$, and $\widetilde{\upsilon_i} = V \upsilon_i$ and the reason for that is this transformations can keep the variance of latent scores without any changing ($V(\upsilon_i) = V I_k V' = \Psi$). The factor loadings matrix Λ is chosen to be a lower block triagonal matrix of full rank and positive diagonal elements (Geweke and Zhou, 1996) using free parameters $f = pq - \frac{q(q-1)}{2}$.

4.2.4 Prior Distributions

This section discuses briefly a complete specification of the prior distributions for all parameters included in this application (see chapter 2 for more details). Since the prior distributions of the underlying variables y^* and the latent variables υ are implicitly determined by the prior distributions of all other parameters and the distributional assumptions about ϵ_i and ξ_i, we have to specify prior distributions for the parameter vector $\theta = vec\{\lambda_0, \Lambda, A, \Sigma, \beta, \gamma, \tau\}$. If we assume that the individual parts of the model are stochastically independent, then the prior distribution yields

$$p(\theta) = p(\lambda_0, \Lambda, A).p(\Sigma).p(\tau).p(\beta, \gamma).$$

The following subsections present briefly the prior distributions of the measurement model $p(\lambda_0, \Lambda, A)$, $p(\Sigma)$ and $p(\tau)$ and of the structural model $p(\beta, \gamma)$.

Prior Distribution of Measurement Model

Prior distribution of intercept, factor loading and direct effects.

Regarding the intercepts factor loadings and direct effects we define a $p.(1+q+d)$ dimensional vector $\overline{\Lambda}$ which contains all parameters of λ_0, Λ and A arranged $\overline{\Lambda} := (\Lambda_{10}, \Lambda_{11}, a_{11}, ..a_{1d}, .., \lambda_{p0}, \lambda_{p1}, .., \lambda_{pq}, a_{p1}, .., a_{pd})$. The prior distribution selected for λ is a $p.(1+q+d)$ dimensional multivariate normal density with the mean $\overline{\lambda}^*$ and the precision matrix $\overline{\Lambda}$ which are chosen according to prior information, i.e.

$$\overline{\lambda} \sim N(\overline{\lambda}^*, \overline{\Lambda}^{*-1})$$
$$p(\overline{\lambda}) \propto constant.$$

We choose noninformative priors for the intercepts λ_0 and the regression coefficients A of direct effects (see Fahrmeir and Raach, 2006). The conjugate prior distribution of the vector of regression coefficients γ_r is a m-dimensional multivariate normal density with the mean γ_r^* and the precision matrix Γ_r^* , i. e. $\gamma_r \sim N(\gamma_r^*, \Gamma_r^{*-1})$. In our analysis, we always choose noninformative priors for all regression parameter γ_r, hence all values of Γ_r^* are set to zero.

Prior Distribution of Structural Model

Prior distribution for Smoothing functions

A prior for smoothing functions $f_{r1}, .., f_{rg}$ is based on a Bayesian P-spline approach (Eilers and Marx (2004))(see chapter 2).

Prior distribution for spatial effect

As mentioned in chapter 2, the prior of spatial effect is based on Markov random filed (Besag, 1974; Besag and Kooperberg, 1995) (see chapter 2).

4.2.5 Fully Posterior Inference

A vector of parameters can be estimated after all parameters are arranged in the parameter vector θ.

$$\theta = vec\{\lambda_0, \Lambda, A, \Sigma, \beta, \gamma, t\}.$$

Hence the posterior distribution

$$p(\theta|y, w, x, u) \propto p(\theta).p(y|\theta, w, x, u).$$

The complete parameter vector is obtained by adding the underlying variables and latent variables to the parameter vector θ leading to the posterior distribution

$$p(\theta, y^*, z|y, w, x, u) \propto p(\theta)p(y, y^*|\theta, w, x, u)$$

Posterior distribution is estimated through MCMC algorithms. Furthermore, there are three different MCMC algorithms can be used which essentially differ in the way of estimating the cutpoints in the case of ordinal indicators. See (Raach, 2005) and (Fahrmeir and Raach, 2006).

Chapter 5

Analysis of Childhood Disease with Geoadditive Probit and Latent Variable Models

Abstract

In this chapter we investigate the impact of various bio-demographic and socio-economic variables on childhood disease with flexible geaodditive probit models. These models allow us to analyze usual linear effects of covariates, nonlinear effects of continuous covariates, and small-area regional effects within a unified, semi-parametric Bayesian framework for modelling and inference. As a first step we employ separate geoadditive probit models (instead of the logit models used in ch.3) to the binary target variables for diarrhea, cough and fever using covariate information from the DHS. Based on these results, we then apply recently developed geoadditive latent variable models where the three observable disease variables are taken as indicators for the latent individual variable "health status" or "frailty" of a child. This modelling approach allows to study the common influence of risk factors on individual frailties of children, thereby automatically accounting

for association between diseases as indicators for health status. We use the probit models in this chapter instead of the logit models which are used in the previous chapter in order to be able to compare the results of the separate geoadditive probit models with the results of the latent variable models (LVM).

5.1 Introduction

The main objective of this chapter is to examine the impact of the socio-economic and bio-demographic factors on childhood disease, including geographical effects as a surrogate for unobserved covariates with spatial information. In our case study, we focus on the analysis for childhood disease in Egypt and Nigeria, using data from the 2003 Demographic and Health Survey (DHS). We will model the impact of various socio-economic, public health and geographical variables on disease of young children in these countries. Statistical analysis will be based on modern Bayesian approaches (as in chapter 3). As a first step, we analyze the impact of various risk factors on the three diseases diarrhea, cough and fever through separate geoadditive probit (instead of logit) models developed in Fahrmeir and Lang (2001) and Brezger and Lang (2005). As a second step, we use geoadditive latent variable models, recently suggested by Raach (2005) and Fahrmeir and Raach (2006). In geoadditive probit latent variable models, the three observable binary disease variables are taken as indicators for the latent individual variable "health status" or "frailty" of a child. This approach is used in order to study the influence of risk factors on individual frailties of children, thereby automatically accounting for association between diseases as indicators for health status. Compared to previous results, our approach can provide new insight to childhood morbidity and mortality in developing countries in general and, more specifically, in Egypt and Nigeria.

Previous studies on child disease have focused on various-socio-economic, demographic or health factors available in specific data sets. Most of these studies, however, have neglected some aspects of spatial effects, see for instance Miller and Hirschhorn (1995), and Miller et al. (1994). Previous work on child disease in Egypt was restricted to few selected or specific towns and governorates. For such work, see Langsten and Hill (1994). Our case study is different from these previous works with respect to the following aspects: first, the analysis studies of spatial differentials of child disease at a highly disaggregated governorates level using a Bayesian approach for geoadditive models. Second, this allows the incorporation of covariate effects in a flexible semi-parametric way, which is not possible through the usual parametric approaches considered in previous works. Third, a latent variable model (LVM) for health status based on binary disease indicators permits modelling of covariates effects on the latent variable through a flexible geoadditive predictor. All computations have been carried out with BayesX-version 1.40 (Brezger, Kneib and Lang, 2005), and R Programs using the MCMC package see Raach (2005) and Fahrmeir and Raach (2006). The rest of this chapter is organized as follows: Section 2 describes geoadditive models and latent variable models, while section 3 contains data analysis, results and discussion for child disease with separate geoadditive models in Egypt and Nigeria. Analyses with latent variables models and comments are given in section 4.

5.2 Bayesian Geoadditive Regression and Latent Variable Models

Geoadditive regression models extend (generalized) linear models for various types of response variables by adding nonparametric terms for nonlinear effects of continuous covariates and geographical effects of a spatial variable to the usual linear part of the predictor. Similarly, predictors in latent variable models can be extended to geoadditive predictors. In the following, we focus on probit models for binary responses, but in general the approach also covers models with continuous, ordered categorical and count variables as observed responses.

5.2.1 Geoadditive Probit Regression

Let $y_1, ..., y_p$ denote p observable binary responses, such as the three disease indicators in our case study, and $x_1, .., x_p$ corresponding covariate vectors. Note that some or even all components of these covariate vectors may be identical, thereby inducing association between the responses. Separate probit models with linear predictors can be defined through

$$P(y_j = 1|x_j) = \Phi(\beta_{0j} + x_j'\beta_j) \quad j = 1, .., p, \tag{5.1}$$

where Φ is the standard normal distribution function. Probit models can be based on Gaussian linear models

$$\widetilde{y_j} = \beta_{0j} + x_j'\beta_j + \epsilon_j, \qquad \epsilon_j \sim N(0,1) \tag{5.2}$$

for unobservable auxiliary variables $\tilde{y}_j$ through the threshold mechanism

$$y_j = 1 \Leftrightarrow \tilde{y}_j > 0,\ y_j = 0 \Leftrightarrow \tilde{y}_j \leq 0. \tag{5.3}$$

Geoadditive probit models are obtained by extending the linear predictor $\eta_j^{lin} = \beta_{0j} + x_j'\beta_j$ to a geoadditive predictor

$$\eta_j^{geo} = \beta_{0j} + x_j'\beta_j + f_1^j(z_1) + .. + f_k^j(z_k) + f_{geo}^j(s).$$

The smooth functions $f_1^j, ..., f_k^j$ represent nonlinear effects of continuous covariates $z_1, ..., z_k$. For simplicity, we only considered the case that these covariates are the same for each predictor $\eta_j^{geo}, j = 1, .., p$. The function f_{geo}^j represents the geographical effect of a spatial variable $s \in \{1, .., d\}$, indicating regions or districts in a country. The geographical effect $f_{geo}^j(s)$ of region s can be interpreted as a surrogate for unobserved variables with geographical information, incomplete or not covered by observable covariates.

It may be split up into a structured part f_{str} for correlated spatial effects, and an unstructured part f_{unstr} for uncorrelated, local spatial effects, see section 3.3. Given the data $(y_{ij}, x_{ij}, z_{i1}, ..., z_{ik}, s_i), i = 1, .., n$, where s_i is the region $\in \{1, .., d\}$ where individual i lives, geoadditive probit models for observations are given by

$$P(y_{ij} = 1|\eta_{ij}^{geo}) = \Phi(\eta_{ij}^{geo}), \quad i = 1, .., n, \ j = 1, .., p \tag{5.4}$$

$$\eta_{ij}^{geo} = \beta_{0j} + x_{ij}'\beta_j + f_1^j(z_{i1}) + .. + f_k^j(z_{ik}) + f_{geo}^j(s_i).$$

Correspondingly, unobservable geoadditive Gaussian models for the auxiliary variables $\tilde{y_j}$ are given by

$$\tilde{y_{ij}} = \eta_{ij}^{geo} + \epsilon_{ij}, \quad \epsilon_{ij} \quad i.i.d \sim N(0, 1). \tag{5.5}$$

The unknown parameters β_{0j}, β_j and functions $f_1^j, ..., f_k^j, f_{geo}^j$ have to be estimated from the data. We follow a semiparametric Bayesian approach as developed in Fahrmeir and Lang (2001) and Brezger and Lang (2005). We assume diffuse, non-informative priors based on Markov Chain Monte Carlo (MCMC) techniques $p(\beta_{0j}) \propto const$, $p(\beta_j) \propto const$. Functions $f_1, ..., f_k$ follow P-spline priors, and the geographical effect f_{geo} is modelled through a Markov random field. Details about these priors are outlined in chapter 2 (section 2.3) of the current work, and the MCMC inference is implemented in BayesX.

5.2.2 Latent Variable Models for Binary Responses

A drawback of separate probit models for each of the binary responses y_j introduced so far is that association among $y_1, ..., y_p$ can only be captured by joint covariates. Latent variable models, as introduced in this section, automatically induce correlation among the responses.

The basic idea of factor analysis and latent variable models (LVM) is that the vector of the p observable variables can be represented, at least partly, by one or more latent factors or variables υ with a lower dimension. As in our case study, where we introduce the latent variable υ "health status" we only consider a one-dimensional latent variable for simplicity. Extension to multi-dimensional latent variables and models with different types of observable responses are presented later in ch.7 (see also Raach (2005) and Fahrmeir and Raach (2006)). The simplest LVM for Gaussian responses $\widetilde{y_j}$, $j = 1, .., p$, and scalar υ is given through

$$\widetilde{y_{ij}} = \lambda_j \upsilon_i + \epsilon_{ij}, \quad i = 1, .., n, \quad j = 1, .., p, \tag{5.6}$$

with i.i.d, Gaussian errors ϵ_{ij}. In this model, υ_i is the unobservable value of individuum i, λ_j is the "factor loading," and $\lambda_j \upsilon_i$ is the effect of υ_i. The restriction to σ_υ=var(υ) $= 1$ is necessary for identifiability reasons; otherwise λ_j would only be identifiable up to the constant $\sigma_\upsilon \neq 1$. If the $\widetilde{y_{ij}}$ cannot be observed directly but only binary indicators

$$y_{ij} = 1 \Leftrightarrow \tilde{y_{ij}} > 0,$$

then we obtain a probit LVM

$$P(y_{ij} = 1|\upsilon_i) = \Phi(\lambda_j \upsilon_i) \quad i = 1, .., n, \quad j = 1, .., p. \tag{5.7}$$

One aspect of the latent variable is that it captures part of the variability of the responses. Secondly, although responses $\widetilde{y_{ij}}$ or y_{ij} are conditionally independent for the given υ_i, they are correlated marginally. These simple models can be extended to geoadditive probit LVMs as follows:

In the most general form, we augment the geoadditive predictors η_{ij}^{geo} in model (5.4) or (5.5) to

$$\eta_{ij}^{geo} + \lambda_j \upsilon_i, \quad i = 1, .., n, \quad j = 1, .., p, \tag{5.8}$$

resulting in the *measurement model*

$$\widetilde{y_{ij}} = \eta_{ij}^{geo} + \lambda_j \upsilon_i + \epsilon_{ij}, \tag{5.9}$$

with i.i.d errors $\epsilon_{ij} \sim N(0,1)$ for the auxiliary variables $\widetilde{y_{ij}}$ and in

$$P(y_{ij}|\eta_{ij}^{geo}, \upsilon_i) = \Phi(\eta_{ij}^{geo} + \lambda_j \upsilon_i), j = 1, .., p$$

for the binary responses.

Secondly, we allow that the latent variable υ is influenced by covariates in form of a *geoadditive structural model*

$$\upsilon_i = u_i'\alpha + f_1(w_{i1}) + ... + f(w_{iq}) + f_{geo}(s_i) + \delta_i, \tag{5.10}$$

with i.i.d. Gaussian errors $\delta_i \sim N(0,1)$. For identifiability reasons as mentioned before it is assumed that $(\delta_i) = 1$, and that the predictor for υ contains no intercept term. The additional covariates $u, w_1, .., w_k$ and the location variable s act directly on the latent variable υ, but indirectly on the observable responses. Covariates included in the structural model (5.10) must not be included in the measurement model at the same time, again for identifiability reasons. In particular, a spatial effect f_{geo} has to be included in either the measurement or the structural model. As for our application, we will restrict the attention to probit LVMs with linear predictors for the measurement model, i.e.,

$$P(y_{ij}|x_{ij}) = \Phi(\beta_{0j} + a_{ij}'\beta_j + \lambda_j \upsilon_i)$$

and geoadditive structural models (5.10) for υ_i. The covariates a_j are different from the covariates $u, w_1, .., w_k$, and they have direct effects β_j on the observed responses. The effects α of u, and the nonparametric effects as well as the spatial effect are indirect effects. We used a_j (instead of x_j) as direct covariates in the case of latent variable model for simplicity.

5.2.3 Priors and Bayesian Inference

To complete the Bayesian model specifications, priors have to be assigned. For the direct effects β_{0j}, β_j and the indirect parametric effects α, we assume diffuse priors

$$p(\beta_{0j}) \propto const, \quad p(\beta_j) \propto const, \quad p(\alpha) \propto const.$$

For the factor loadings, we specify informative Gaussian priors

$$p(\lambda_j) \propto N(0, \sigma_j^2),$$

with $\sigma_j^2 = 1$ as the standard choice, to avoid the so-called Heywood cases (see e.g. Raach, 2005).

Priors for functions

For a function $f(w)$ of a continuous covariate w, we assume Bayesian P-spline-priors as in Brezger and Lang (2005).

Priors for spatial effects

We usually split f_{geo} (s) into a smooth structured effect and an unstructured effect, i.e.

$$f_{geo}(s) = f_{str}(s) + f_{unstr}(s)$$

where $f_{str}(s)$ models global spatial trends while $f_{unstr}(s)$ captures local effects. For $f_{str}(s)$ we assume a Markov random field prior (Beseg, York and Mallie, 1990), while the unstructured effects are i.i.d. random variables. For more details, see also chapter 2.

5.3 Statistical Analyses and Results

Statistical analyses were performed in two steps:

First, we fitted separate geoadditive probit models to the following three diseases: diarrhea, fever and cough. A main purpose of this step was model selection, to model effects of the continuous covariates, and to see if there are sizeable spatial effects. Based on preliminary exploratory analyses not shown here, we used the Deviance Information Carterion (DIC) of Spiegelhalter et al.(2002) to select models in a formal way. Section 4.1 presents results of this first data analysis step. In the second step, we then applied geoadditive probit LVMs to analyze the data. While the DIC is now commonly accepted as a standard tool for selecting probit or logit models, its performance for LVM model choice is not yet well understood. It was decided to choose the covariates used in equation (5.9) for the measurement model, which have direct effects on the disease indicators; or in the case of the structural equation (5.10), those have indirect effects via their common impact on the latent variable "health status," we therefore proceeded more informally: if the effects of covariates turned out to be significantly different (in terms of confidence intervals) for the three diseases, we decided to keep them in the measurement model, otherwise covariates were included in the geoadditive predictor of the structural equation for the latent variable. The results are presented in section 4.2.

5.3.1 Analyses with Separate Geoadditive Models

We present results for the following probit models, selected from a longer hierarchy of models. The responses y_j, $j = 1$ (diarrhea) , 2 (fever), 3 (cough) are coded as

$$y_i = \begin{cases} 1 & : \text{if child had disease 2 weeks prior to the survey} \\ 0 & \text{if not} \end{cases} \tag{5.11}$$

The following covariates were considered in the analysis in both countries:

Metrical covariates

Chage: Child's age in months.

BMI: Mother's body mass index.

Mageb: Mother's age at birth.

Categorical covariates (in effect coding)

male: Child's sex: male or female (reference category).

educ: Mother's educational attainment: incomplete primary, complete primary, and incomplete secondary school or complete secondary school and higher eduction (reference category).

trepr: Whether mother had treatment during pregnancy: yes or no (reference category).

anvis: Whether mother had antenatal care: yes or no (reference).

water: Source of drinking water: controlled water or no (reference category).

toilet: Has flush toilet at household: yes or no (reference category).

urban: Locality where respondent lives: urban or rural (reference category).

radio: Has a radio at household: yes or no (reference category).

elect: Has electricity: yes or no (reference category).

work: Mother's current working status: Working or not (reference).

Spatial covariate

reg: Governorate or region where respondent resides.

The predictors of the models considered in this section are as follows:

M0: Included only district-specific effects.

$$M0 : \eta_{ij} = \beta_0 + f_{str}(reg) + f_{unstr}(reg) \tag{5.12}$$

M1: Includes all categorical covariates and the metrical covariates.

$$M1 : \eta_{ij} = \beta_{0j} + f_j(Chage) + f_j(BMI) + f_j(Mageb) + w_i'\gamma_j \tag{5.13}$$

M2: Adds district-specific effects to Model 1.

$$M2 : \eta_{ij} = \beta_{0j} + f_j(Chage) + f_j(BMI) + f_j(Mageb) + f_{str}(reg) + f_{unstr}(reg) + w_i'\gamma_j \quad (5.14)$$

$$M3 : \eta_{ij} = \beta_{0j} + f_j(Chage) + f_j(BMI) + f_j(Mageb) + f_{str}(reg) + f_{unstr}(reg) + z_i'\gamma_j \quad (5.15)$$

In these models, β_0 is a constant term and the covariate vector w in models M1 and M2 contains all the bio-demographic and health factors. In model M3 the vector w is reduced to the vector z by omitting factors of education, type of toilet and source of water. The metrical covariates child's age, mother's BMI and mother's age at birth are allowed to have a non-linear effect on the diseases of child as well as the spatial effects f_{str} and f_{unstr}. It turned out that model M3 for each type of diseases is superior in terms of the DIC.

Results

In the preliminary analysis, we aim to separate the two kinds of spatial effects included in model M0 to estimate a structured and an unstructured effect. In a further step, we include the categorical covariates and the metrical covariates in the analysis as shown in models M1, M2 and M3. The results for these models are given in tables 5.2 through 5.19 for the categorical covariates, in figures 6-8 for the effects of the continuous covariates of child's age, mother's BMI and mother's age at birth, and in figures 5.1, 5.8, 5.3, 5.10, 5.5 and 5.12, which suggest district variation in the prevalence of diarrhea, cough, and fever in Egypt and Nigeria, respectively.

Diarrhea

Tables 5.2 through 5.7 display the estimated categorical effects of these variables (male, urban, mother working status, mother had treatment during pregnancy, antenatal visit, availability of radio, availability of electricity, source of drinkable water, mother's education, and toilet facility) on diarrhea disease in both countries. The results of Egypt indicate a significant impact of sex (male), locality of residence, antenatal visit, having radio (only in

M1) and mother had treatment during pregnancy on disease of diarrhea and a significant impact of mother had treatment during pregnancy, antenatal visit, and having radio (only in M1) in Nigeria. However, antenatal visit has a negative impact on diarrhea disease in Nigeria and a positive effect in the case of Egypt. This analysis also suggests that mother's education, mother working status, toilet facility, availability of electricity and source of drinkable water have little or non significant effects in both countries.

With regard to the non-linear effects, figure 5.2 and 5.8 show from top to bottom: the (nonlinear) effects of age of the child, mother's body mass index and mother's age at birth for models M1, M2 and M3, respectively, modelled through Bayesian P-splines. In Egypt, the nonlinear effect of child's age suggested that there is continuous and serious worsening of children's health status up to about 11 months of age, with an almost linear decline thereafter. In Nigeria, the nonlinear impact of child's age also suggested a high risk of getting diarrhea during the first 11 months of age, but the impact goes to be almost linear until about 25 months of age, with linear decline thereafter. The top to third (from the top) of the right panel indicate that the impact of a mother's BMI on diarrhea is only slight. There is some evidence that the children of mothers whose have a BMI less than 25 face a lower risk of disease (even though there are few mothers with BMI between 15 and 20). For BMI larger than 43-45, there are few observations and the credible intervals gets wider. A somewhat higher risk for diarrhea seems to exist for mothers who have a BMI between 27 and 30, where a bump appears. On the other hand, the impact of mother's BMI on diarrhea in Nigeria (right panel of figure 5.8) is slight with almost linear for mothers with BMI up to about 30 and the impact seem almost linear decline thereafter. In addition, we find the influence of mother's age (second panel from the bottom to the top bottom of figure 5.2) on diarrhea in Egypt seems to be in the form of an inverse U-shape. It shows that the mother's age has a slight impact on diarrhea, however the children from mothers who are in age group (18-22 years) are at a higher risk of diarrhea compared to children from mothers in other age groups. Further, the pattern of mother's age (second panel from the bottom to the top bottom of figure 5.8) in Nigeria is very similar to that of Egypt's.

With regard to spatial effects, figures 5.1 and 5.7 display the estimates of the spatial effect (the levels correspond to "high risk of morbidity" (green colored) and "low risk" (red colored) for Egypt and Nigeria, respectively. The colored maps show posterior means of structured random effects on diarrhea (right panels) and its corresponding posterior mean of unstructured random effects (left panels). For the model M0, model M2 and model M3 for the diarrhea disease, the geographical pattern of regions in the right panel of figures 5.1 and 5.7 reflects the estimated posterior means of the structured random effects on diarrhea. Obviously, there exists a district-specific geographical variation in the level of the disease in Egypt (figure 5.1) based on the 2003 EDHS. The pattern reveals that significant high rates of illness are associated with the Upper Egypt area (Minya, Amarna, Luxor, Esna, Edfu, Aswan, ...), some cities and rural areas in the Nile Delta and in Eastern Cairo (Sinai). Upper Egypt implies a relative higher risk of having a diarrhea disease and knowing the characteristics of the region, the result is not surprising (see chapter 3 of the current work, discussion). The left panel also reveals a higher risk of diarrhea morbidity in the upper area in spite of being surrounded by some districts with lower risk. According to spatial effect in Nigeria, illness rates are significantly high in Borono, Adanowa, Taraba (northeastern regions through southeastern part), while Bauchi (central region) have substantially lower significant spatial effects. Non significant effects are observed in other states.

Model (country)	Deviance	pD	DIC
	Diarrhea		
M0 (Egypt)	6364.46	15.45	6395.38
M1 (Egypt)	5433.27	36.53	5506.34
M2 (Egypt)	5432.74	36.91	5506.55
M3 (Egypt)	5311.83	46.50	5404.84
M0 (Nigeria)	4419.83	28.58	4477.01
M1 (Nigeria)	3152.81	28.17	3209.16
M2 (Nigeria)	2921.11	52.68	3026.48
M3 (Nigeria)	2923.32	51.39	3026.10
	Fever		
M0 (Egypt)	7892.49	12.98	7918.47
M1 (Egypt)	6972.51	36.74	7045.99
M2 (Egypt)	6904.23	48.043	7000.32
M3 (Egypt)	6911.25	44.38	7000.02
M0 (Nigeria)	6079.09	28.82	6136.74
M1 (Nigeria)	3969.13	29.26	4027.66
M2 (Nigeria)	3826.83	52.26	3931.36
M3 (Nigeria)	3826.09	49.68	3930.47
	Cough		
M0 (Egypt)	7076.87	14.43	7106.38
M1 (Egypt)	6432.95	35.92	6504.78
M2 (Egypt)	6330.83	48.96	6428.75
M3 (Egypt)	6336.81	45.15	6427.11
M0 (Nigeria)	5312.56	30.58	5373.74
M1 (Nigeria)	3596.00	29.64	3655.29
M2 (Nigeria)	3396.35	58.32	3512.99
M3 (Nigeria)	3398.02	55.83	3509.69

Table 5.1: The Deviance Information Criterion (DIC)

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.03^*	0.166	-1.398	-1.024	-0.741
male	0.057^*	0.020	0.018	0.057	0.094
urban	-0.067^*	0.023	-0.112	-0.066	-0.021
work	0.015	0.0267	-0.036	0.014	0.071
trepr	0.069^*	0.031	0.006	0.070	0.128
anvis	0.079^*	0.021	0.041	0.079	0.122
radio	-0.059^*	0.025	-0.108	-0.059	-0.011
elect	0.024	0.093	-0.151	0.022	0.207
water	0.006	0.028	-0.047	0.007	0.058
educ	-0.030	0.024	-0.079	-0.032	0.016
toilet	-0.041	0.047	-0.139	-0.042	0.053

Table 5.2: Fixed effects (M1) on diarrhea-Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.093^*	0.159	-1.415	-1.087	-0.799
male	0.058^*	0.020	0.017	0.059	0.098
urban	-0.062^*	0.024	-0.112	-0.063	-0.011
work	0.014	0.026	-0.036	0.015	0.067
trepr	0.064	0.0321	-0.0007	0.064	0.125
anvis	0.089^*	0.021	0.041	0.088	0.133
radio	-0.043	0.026	-0.095	-0.043	0.006
elect	0.009	0.095	-0.170	0.008	0.207
water	0.016	0.029	-0.036	0.015	0.075
educ	-0.032	0.024	-0.077	-0.032	0.019
toilet	-0.053	0.045	-0.145	-0.054	0.034

Table 5.3: Fixed effects (M2) on diarrhea-Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.116^{*}	0.169	-1.469	-1.119	-0.806
male	0.060^{*}	0.019	0.021	0.060	0.097
urban	-0.062^{*}	0.024	-0.109	-0.063	-0.016
work	0.010	0.025	-0.042	0.010	0.057
trepr	0.065^{*}	0.031	0.002	0.065	0.129
anvis	0.080^{*}	0.022	0.036	0.079	0.123
radio	-0.051	0.025	-0.101	-0.050	0.001
elect	-0.002	0.094	-0.177	-0.003	0.201

Table 5.4: Fixed effects of model(M3) on diarrhea-Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.29^{*}	0.139	-1.594	-1.277	-1.033
male	0.047	0.0269	-0.004	0.045	0.103
urban	0.006	0.033	-0.057	0.005	0.069
work	-0.014	0.028	-0.070	-0.012	0.039
trepr	0.075^{*}	0.034	0.011	0.074	0.144
anvis	-0.107^{*}	0.028	-0.160	-0.107	-0.057
radio	-0.068^{*}	0.028	-0.123	-0.069	-0.010
elect	-0.021	0.030	-0.08	-0.022	0.041
water	-0.038	0.038	-0.108	-0.038	0.042
educ	-0.026	0.047	-0.125	-0.025	0.063
toilet	-0.094	0.052	-0.201	-0.093	0.007

Table 5.5: Fixed effects (M1) on diarrhea-Nigeria.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.385^{*}	0.174	-1.737	-1.372	-1.048
male	0.047	0.025	-0.003	0.047	0.097
urban	-0.039	0.035	-0.107	-0.039	0.032
work	0.011	0.030	-0.045	0.011	0.069
trepr	0.033	0.036	-0.039	0.033	0.105
anvis	-0.059	0.035	-0.134	-0.061	0.008
radio	-0.041	0.031	-0.106	-0.040	0.019
elect	0.026	0.034	-0.036	0.027	0.094
water	-0.050	0.044	-0.130	-0.052	0.038
educ	-0.0126	0.048	-0.110	-0.010	0.081
toilet	-0.038	0.055	-0.152	-0.038	0.070

Table 5.6: Fixed effects (M2) on diarrhea-Nigeria.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-1.32^{*}	0.165	-1.65	-1.32	-1.012
male	0.047	0.026	-0.002	0.046	0.101
urban	-0.052	0.034	-0.119	-0.054	0.018
work	0.0145	0.028	-0.040	0.014	0.069
trepr	0.035	0.034	-0.032	0.035	0.100
anvis	-0.064^{*}	0.033	-0.128	-0.065	-0.0009
radio	-0.039	0.032	-0.103	-0.039	0.025
elect	0.0168	0.033	-0.046	0.0159	0.079

Table 5.7: Fixed effects of model (M3) on diarrhea-Nigeria.

Fever

The fixed parameters show that the prevalence of fever in Egypt (tables 5.8 through 5.10) is higher among infants from mothers who are working, males, and children from mothers who obtained antenatal visits during pregnancy. Availability of radio in the household is associated with a lower risk of fever morbidity. On the other hand, the results suggest that mother's educational attainment, whether the mother received injection during pregnancy or not, availability of a flush toilet, availability of electricity, source of drinkable water and locality of residence have only a slight influence on fever morbidity in Egypt. In Nigeria, the results suggest that the prevalence of fever (tables 5.11 through 5.13) is low among children who live in urban areas, have a flush toilet in the household, but children from mothers who obtained treatment during pregnancy are at a higher risk of fever. However, urban is only significant in M1. In addition, mother who obtained antenatal care during pregnancy, had access to electricity and radio have a lower significant effect on fever, but source of drinkable water, mother's educational attainment and sex of children have non significant influence on fever morbidity in Nigeria.

Figures 5.4 and 5.10 show the nonlinear effects of a child's age on fever for model M1 (top left), model M2 (second left) and model M3 (top third) in both countries, respectively. The impact of a child's age is quite similar in the three models in Egypt and Nigeria as well. They show that deterioration sets in right after birth and continues, up to 11-12 months, but then the age effect declines more or less steadily until 25-26 months. In Nigeria, however, it is apparent that a higher risk for fever comes into view for children who are in age group 27-30 as seen in figure 5.10 (top left through third panel from top). The effect of mother's BMI on fever is shown in figures 5.4 and 5.10 (top right through third panel from top). It is observed that mother's BMI has a slight significant impact on child health status in both countries. Furthermore, it declines for mothers with a BMI of less than 20, and is less pronounced for mothers with BMI between 20-35 in both countries, in spite of a blip between BMI of 30 and 35, which is caused by overweight mothers in Egypt, and over a BMI of 40, there are only few observations (wide credible interval). Unexpectedly, the effect of mother's BMI $f(BMI)$ in the three models turns out to be almost linear for both countries.

With regard to the non-linear effect of mother's age at birth on fever morbidity, the fourth left panel from the top of figures 5.4 and 5.10 displays that children from younger mothers (< 20 years) are at considerably higher risk of morbidity compared to children from mothers who are in the middle-aged group (25-35) and the impact of mother's age on fever disease is quite similar for both countries.

The overall pattern is very similar to diarrhea's.

The geographical pattern of district-specific effects for fever in figure 5.3 indicates that significant high illness rates are associated with the Egyptian governorates Suez, El Arish, Ismalia and Sinia "in the southwestern area". There is a variation in the level of illness rates of children in Egypt, and this variation could be attributed to environmental risks, which in turn influence exposure to disease. The unstructured effects are similar to the structured effects. The gray area, however, indicates that no children live there.

The spatial effect in Nigeria (figure 5.9) indicates that highly significant rates of fever illness are associated with northeastern parts of Nigeria. High prevalence is noticeable in Adanowa state. In the southeastern regions, significant high fever rates are observed in Taraba, Plateatu and Bauchi states (see chapter 3, section 3.5).

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.400^*	0.120	-0.648	-0.396	-0.152
male	0.044^*	0.017	0.009	0.043	0.080
urban	0.013	0.020	-0.024	0.013	0.052
work	0.052^*	0.021	0.008	0.053	0.097
trepr	0.025	0.029	-0.038	0.027	0.083
anvis	0.080^*	0.019	0.040	0.080	0.118
radio	-0.064^*	0.023	-0.111	-0.064	-0.016
elect	-0.110	0.086	-0.286	-0.110	0.073
water	0.027	0.027	-0.024	0.026	0.082
educ	-0.027	0.020	-0.067	-0.026	0.013
toilet	-0.031	0.044	-0.120	-0.032	0.057

Table 5.8: Fixed effects of model (M1) on fever-Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.373^*	0.130	-0.626	-0.384	-0.110
male	0.043^*	0.0168	0.012	0.043	0.077
urban	0.007	0.022	-0.037	0.0066	0.049
work	0.050^*	0.023	0.005	0.050	0.094
trepr	0.029	0.029	-0.031	0.029	0.087
anvis	0.080^*	0.020	0.039	0.080	0.118
radio	-0.064^*	0.022	-0.107	-0.064	-0.019
elect	-0.103	0.081	-0.264	-0.102	0.048
water	0.034	0.026	-0.018	0.034	0.085
educ	-0.029	0.021	-0.0718	-0.030	0.0123
toilet	-0.051	0.0413	-0.133	-0.054	0.032

Table 5.9: Fixed effects of model (M2) on fever-Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.270^{*}	0.186	-0.603	-0.275	0.095
male	0.046^{*}	0.017	0.009	0.044	0.080
urban	0.006	0.021	-0.036	0.006	0.049
work	0.043	0.023	-0.003	0.045	0.089
trepr	0.025	0.030	-0.036	0.025	0.0868
anvis	0.075^{*}	0.0198	0.039	0.074	0.115
radio	-0.069^{*}	0.024	-0.119	-0.068	-0.019
elect	-0.211	0.167	-0.536	-0.200	0.104

Table 5.10: Fixed effects of model (M3) on fever-Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.641^{*}	0.127	-0.898	-0.641	-0.384
male	0.008	0.022	-0.032	0.008	0.052
urban	-0.058^{*}	0.028	-0.113	-0.058	-0.002
work	-0.012	0.024	-0.062	-0.0113	0.032
trepr	0.081^{*}	0.030	0.021	0.080	0.145
anvis	-0.008	0.029	-0.068	-0.009	0.049
radio	-0.029	0.028	-0.089	-0.030	0.025
elect	-0.033^{*}	0.027	-0.089	-0.032	0.022
water	0.040	0.032	-0.018	0.039	0.109
educ	0.022	0.041	-0.0625	0.022	0.101
toilet	-0.168^{*}	0.045	-0.254	-0.167	-0.079

Table 5.11: Fixed effects of model (M1) on fever-Nigeria.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.639^{*}	0.133	-0.901	-0.639	-0.366
male	0.018	0.024	-0.030	0.018	0.065
urban	-0.050	0.031	-0.107	-0.052	0.013
work	0.024	0.027	-0.027	0.026	0.075
trepr	0.073^{*}	0.031	0.010	0.072	0.139
anvis	0.022	0.031	-0.037	0.022	0.082
radio	-0.025	0.029	-0.082	-0.026	0.034
elect	-0.012	0.030	-0.076	-0.012	0.047
water	0.024	0.034	-0.045	0.025	0.090
educ	0.019	0.041	-0.060	0.019	0.101
toilet	-0.117^{*}	0.048	-0.214	-0.116	-0.019

Table 5.12: Fixed effects of model (M2) on fever-Nigeria.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.600^{*}	0.125	-0.844	-0.605	-0.363
male	0.017	0.023	-0.028	0.017	0.064
urban	-0.048	0.030	-0.106	-0.047	0.010
work	0.025	0.026	-0.025	0.025	0.079
trepr	0.071^{*}	0.030	0.010	0.072	0.130
anvis	0.018	0.030	-0.041	0.016	0.080
radio	-0.029	0.031	-0.092	-0.029	0.026
elect	-0.019	0.031	-0.075	-0.020	0.0418

Table 5.13: Fixed effects of model (M3) on fever-Nigeria.

Cough

The results indicate that children from mothers who attended an antenatal care during pregnancy, and currently working face a high rate of cough disease compared to children from mothers who are not working and did not attended any care. The results also suggested that ownership of radio facility has a negative impact on cough disease in Egypt. It is observed that the boys under 5 years are more likely to get cough morbidity than girls. The rest of categorical covariates have either a negligible impact or an insignificant effect on cough morbidity (tables 5.14 through 5.16). In Nigeria, the results (tables 5.17 through 5.19) observed that only the covariate of whether the mother had treatment during pregnancy or not has a significant effect on cough disease overall for the three models. Further, the results indicate that some covariates such as availability of electricity, source of water, place of residence, and education attainment are only at the borderline to significance.

The non-linear effect of child's age for model M1 (left top panel of figures 5.6 and 5.12), model M2 (second left from top) and model M3 (third left from top) has a similar pattern to diarrhea and fever. The same is true for mother's BMI and mother's age at birth, for both countries.

Spatial effect on cough in Egypt is seen in figure 5.5. The results suggest that significantly high rates of cough illness are associated with Damietta, Dakhalia and Esmaliyia.

The results of spatial effect, which are shown in figure 5.11, indicate that the northeastern part of Nigeria and some states in southern parts of the country, such as Cross River, Bayclsa, Gombe, and Yobe are associated with high presence of cough disease.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.483^{*}	0.192	-0.868	-0.477	-0.095
male	0.044^{*}	0.018	0.007	0.044	0.080
urban	0.050	0.021	0.006	0.050	0.090
work	0.070^{*}	0.023	0.025	0.070	0.114
trepr	0.011	0.029	-0.047	0.012	0.068
anvis	0.082^{*}	0.020	0.042	0.082	0.124
radio	-0.055^{*}	0.024	-0.104	-0.055	-0.012
elect	-0.072	0.173	-0.399	-0.074	0.272
water	-0.019	0.025	-0.069	-0.019	0.029
educ	-0.027	0.022	-0.071	-0.028	0.017
toilet	-0.035	0.046	-0.116	-0.035	0.065

Table 5.14: Fixed effects (M1) on cough-Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.410^{*}	0.199	-0.795	-0.415	-0.008
male	0.045^{*}	0.0178	0.010	0.045	0.079
urban	0.042	0.022	-0.003	0.044	0.086
work	0.064^{*}	0.0239	0.0158	0.065	0.109
trepr	0.025	0.030	-0.040	0.026	0.084
anvis	0.068^{*}	0.0216	0.027	0.068	0.112
radio	-0.056^{*}	0.024	-0.102	-0.056	-0.009
elect	-0.052	0.175	-0.429	-0.053	0.278
water	-0.019	0.027	-0.073	-0.021	0.039
educ	-0.037	0.0218	-0.079	-0.038	0.005
toilet	-0.054	0.044	-0.144	-0.053	0.035

Table 5.15: Fixed effects model (M2) on cough-Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.420^{*}	0.199	-0.811	-0.420	-0.036
male	0.046^{*}	0.017	0.010	0.046	0.077
urban	0.033	0.021	-0.009	0.032	0.072
work	0.0596^{*}	0.0235	0.010	0.060	0.104
trepr	0.026	0.029	-0.027	0.0273	0.084
anvis	0.059^{*}	0.020	0.019	0.059	0.099
radio	-0.066^{*}	0.024	-0.115	-0.066	-0.017
elect	-0.090	0.177	-0.422	-0.093	0.270

Table 5.16: Fixed effects (M3) on cough-Egypt.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.748^{*}	0.129	-0.996	-0.752	-0.480
male	-0.003	0.025	-0.051	-0.002	0.048
urban	-0.032	0.028	-0.089	-0.031	0.022
work	0.029	0.0248	-0.025	0.029	0.077
trepr	0.141^{*}	0.033	0.076	0.139	0.207
anvis	0.058	0.028	-0.001	0.058	0.116
radio	-0.032	0.029	-0.09	-0.03	0.023
elect	-0.047	0.029	-0.107	-0.048	0.007
water	-0.030	0.036	-0.100	-0.030	0.038
educ	0.009	0.038	-0.070	0.010	0.079
toilet	0.0016	0.046	-0.089	0.00003	0.100

Table 5.17: Fixed effects (M1) on cough- Nigeria.

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.748^*	0.129	-0.996	-0.752	-0.480
male	-0.003	0.025	-0.051	-0.0029	0.048
urban	-0.032	0.028	-0.089	-0.031	0.022
work	0.029	0.024	-0.025	0.029	0.077
trepr	0.141^*	0.033	0.076	0.139	0.207
anvis	0.058	0.028	-0.001	0.058	0.116
radio	-0.032	0.029	-0.094	-0.032	0.023
elect	-0.047	0.029	-0.107	-0.048	0.007
water	-0.030	0.036	-0.100	-0.030	0.038
educ	0.009	0.038	-0.070	0.010	0.079
toilet	0.0016	0.046	-0.089	0.0003	0.106

Table 5.18: Fixed effects model (M2) on cough-Nigeria.

5.3.2 Discussion

Fixed Effects

As for child's gender, it is widely believed that probability of disease is higher for males due to biological reasons. Although, boys are noticeably more likely than girls to be taken to a provider for treatment (EDHS 2003). However, some studies show higher female mortality indicating gender discrimination. The results show that a child's gender is mostly significant and has a large impact on the three types of diseases in Egypt. In Nigeria, this variable is insignificant for three types of diseases.

The effects of urban versus rural place of residence are different for the three diseases: For diarrhea, living in urban areas lowers the risk, for fever and cough the effect is not significant for children from urban vs.rural areas. These results support the important role of the public health policy in rural-urban disparities.

Mothers who attended a clinic to receive antenatal care during the period of pregnancy are expected to have lower problems in comparison to those who had not received any care. The results for Egypt, however, suggest the

Variable	Mean	S.dv	2.5%	median	97.5%
const	-0.791^*	0.143	-1.076	-0.793	-0.513
male	0.0025	0.023	-0.048	0.002	0.049
urban	-0.060	0.031	-0.119	-0.062	0.0031
work	0.015	0.027	-0.039	0.014	0.065
trepr	0.105^*	0.032	0.045	0.105	0.170
anvis	0.033	0.033	-0.030	0.034	0.102
radio	0.0035	0.030	-0.050	0.0012	0.067
elect	0.002	0.031	-0.060	0.0033	0.062

Table 5.19: Fixed effects (M3)on cough-Nigeria.

contrary: the factor antenatal visit has a positive effect on the indicators of disease! A possible reason could be that there are few of mothers who obtained antennal visits frequently during their pregnancy. In addition, there are only 10% who had treatment during their pregnancy or maybe the reason of getting care was not related to their pregnancy. Furthermore, the same reason could exist for the variable *trepr*, which has a positive significant effect on the three types of disease in Nigeria.

The ownership of radio facilitates the acquisition of disease and vaccination information, allowing a more effective allocation of resources to produce child health. Therefore, it has a negative significant effect on the morbidity as suggested by the results for Egypt, but in Nigeria it has nonsignificant effect.

Concerning current working status of mother, these results suggest a significant effect of this variable on fever and cough morbidity in Egypt, however the effect is positive. The problem is when mothers engage in out-of-home employment it curtails the duration of full breastfeeding and necessitates recently introduced supplementary feeding, often by the illiterate care-takers, and that could have a side effect on the health of child in the early months.

Availability of the flush toilet in the household is associated only with a lower risk of fever in Nigeria. The same for availability of electricity with diarrhea in Nigeria.

Non-Linear Effects

In general, the results show that the risk of having diseases in the two-week reference period reaches its peak at 11 months and then begins to fall with increasing age of the child. This pattern resembles those found in many studies of sub-Saharan Africa. The prevalence of disease was found to be highest among children 6-12 months of age, the period when most children are weaned. In addition to breast milk, inborn immunity and less exposure to contaminated agents during the early period also contributes to the lower prevalence of diarrhea. On the other hand, prevalence is quite high when the child has lost inborn immunity and when it is exposed to different types of infections by eating food prepared with contaminated water and from an unhealthy environment.

Likewise, the effect of mother's age at birth is almost linear in Egypt, particularly in the interval age between 20 and 27 years. The curve has a slight bathtub shape, indicating that children from younger mothers (12-20) have higher risk, compared to mothers 20-35 years old. The results reflect a slight effect of mother's age at birth on the morbidity of children. In Nigeria, the impact of mother's age is slight and almost linear.

In the literature, the influence of the body mass index (BMI) of the mother is sometimes expected to be inversely U-shaped. Parents with low BMI values are malnourished and are therefore likely to have undernourished and weak children. At the same time, very high BMI values indicate poor quality of the food and hence, may also imply weakness of the children in our study. The results of Egypt indicate that a mother's BMI of 27-30 greatly increases the effect on child morbidity. Beyond a BMI of 30, the effect remains at a low level equilibrium. The higher impact of BMI through the interval between 27-30, indicates poor quality of food for mothers and hence, may imply malnutrition of the child and affect the health of the child. For Nigeria, it has a slight effect on cough and fever.

Spatial Effects

The Egyptian regions used in this study and in previous studies are metropolitan, Lower Egypt, Upper Egypt and border areas. Ninety-five percent of the population of Egypt lives in the first three regions. The metropolitan governorates essentially comprise the four major cities of Cairo, Alexandria, Port-Said and Suez, all in northern Egypt. Lower Egypt (essentially the region of the Nile Delta) is also in the northern part of Egypt, and Upper Egypt is the area south of Cairo, with governorates largely following the meandering upper parts of the Nile. The border areas are the less populated desert areas bordering the Red Sea, the Sinai, and the vast Marsa Matruh and El Wadi El Gadid areas west of the Nile. Generally, childhood diseases appear to have higher influence on child in the north-east part, affecting the most of districts there. Food insecurity associated with water supplies and quality of water could be a reason for these negative effects in this area.

In Nigeria, there is a sizeable difference between disease in the eastern parts of the country and the significantly better health status in the northern, and central parts. We can see from the results that southeastern regions through some regions in the north part are associated with a high rate of childhood disease. That is because, as suggested by previous studies, is present a high level of pollution due to petroleum production in those regions. For this reason, the pollution in this area affected the health of children through the water pollution that influences access to drinkable water sanitation (see also chapter 3, section 3.5).

5.3.3 Comparison with Previous Results

In this chapter, we explored determinants of child disease in Egypt and Nigeria using geoadditive probit models. Compared to the results using geoadditive logit model in chapter 3, particulary the results included in section 3.6, the results are very similar. Our focus in comparison on the results of M2 with the results have obtained in section 3.5, because the covariates used in both are the same, but the models are different.

The results of fixed effects parameters have shown that the coefficients for the covariates in the logit model are approximately 1.7, the coefficients for the covariates in probit model known by the statistical literature.

Concerning the effects of the covariates and their significant levels, the re-

Parametrer	Mean	Std	2.5%	97.5%
		Factor Loadings		
1. Fever λ_{11}	2.2	0.34	1.78	3.03
2. Cough λ_{21}	0.87	0.04	0.77	0.959
3. Diarrhea λ_{31}	0.67	0.03	0.616	0.73

Table 5.20: Results of Model LVM0 for Egypt with $\eta = 0$.

sults are quite similar. Furthermore, the nonlinear effects of metrical covariates are based on P-splines in this chapter instated of second order random walk which are used in chapter 3, therefore the functions are smoother compared to the functions of the second order random walk . For the spatial effects the results are very similar to the results of chapter 3.

5.4 Analyses with Latent Variable Models

As previously discussed in section 3, we now investigate how the three diseases can be interpreted as indicators of a latent variable υ "health status" of children, how much of the variation of υ can be explained through a geoadditive predictor, and which covariates have a direct effect on the disease indicators. This concept does not only allow us to analyze the impact of covariates on health status, it also automatically introduces a correlation among disease indicators. To demonstrate the latter property, we first consider a classic model without any covariates, i.e. in turns of auxiliary variables.

(LVM0):

$$P(y_{ij} = 1|\upsilon_i) = \Phi(\lambda_j \upsilon_i), \quad \upsilon_i \sim N(0, 1) \tag{5.16}$$

and $\eta = 0$, so that $\upsilon_i \sim N(0, 1)$. Tables 5.20 and 5.21 show the estimates for the factor loadings λ_j, $j = 1, 2, 3$ implying considerable (marginal) correlation.

Our next model is selected on the basis of the separate analyses as explained at the beginning of this section. This leads to the latent variable model

Parametrer	Mean	Std	2.5%	97.5%
Factor Loadings				
1. Fever λ_{11}	1.48	0.12	1.30	1.72
2. Cough λ_{21}	0.99	0.064	0.87	1.12
3. Diarrhea λ_{31}	0.69	0.042	0.61	0.77

Table 5.21: Results of Model LVM0 with $\eta = 0$ for Nigeria.

$$P(y_{ij}|x_{ij}) = \Phi(\beta_{0j} + a'_{ij}\beta_j + \lambda_j \upsilon_i),\ j = 1, 2, 3$$

with the structural model

$$\upsilon_i = u'\alpha + f_1(Chage_i) + f_2(BMI_i) + f_3(Mageb_i) + f_{geo}(reg_i) + \delta_i$$

for the latent variable. The vector a' (measurement model) comprises the covariates with direct effects (such as urban, availability of electricity and controlled water in LM1 for Egypt) on y_j, and u comprises the remaining categorical covariates (such as sex, mother's education, etc. in LM1 for Egypt) having common effects on the latent variable υ. Because the patterns for the nonparametric functions and the spatial effects were rather similar in the separate analyses, they were included in the geoadditive predictor for υ.

The results of latent variable models for categorical covariates are in table 5.22. Factor loadings are slightly lower than for the factor analysis without covariates.

Because indirect effects affect the latent variable, they cover a larger range of values and thus exert more influence on the variability of the indicators, even if the factor loadings are slightly lower.

The results for Egypt show that the parametric indirect covariates of male, antenatal visit, having radio, and mother's working status have a significant effect on the latent variables. The results indicate that the mother's education, ever had treatment during pregnancy and toilet facility have only a non-significant or slight effect on the latent variables. Concerning the categorical direct covariates, the results indicate a significant effect of urban on cough and diarrhea. However, the effect of urban on cough is positive. The reason is that only one instead of three separate effects have to be estimated. The results of LVM1 are quite consistent with the previous results which are obtained using geoadditive models for each kind of disease in section 4.1. The insignificant (access to electricity and water) parametric direct covariates were included in the parametric indirect effects in LVM2 (table 5.23) and they still have nonsignificant impacts on the indicators of health status in Egypt, therefore, we excluded these covariates in model LVM3 (table 5.24).

The pattern for non-linear effects on the latent variable health status closely resembles the patterns of separate analyses. Furthermore, there is no noticeable difference between the nonlinear effects by model LVM2 and model LVM3. Therefore, only the results of model LVM2 are reported here.

The spatial effect is displayed in figure 5.14, and shows that the northeast has an influence on the latent variable associated with high illness rates. These areas face problems with health conditions, level of sanitation and water supplies that could lead to a high level of infections among children under 5 living in these areas.

In Nigeria (table 5.25), the parametric indirect covariate of *trepr* (whether the mother had treatment during pregnancy) has a significant positive effect on the latent variables. With regard to the parametric direct covariates, the results show that urban location has a significant negative effect on an indicator of fever λ_{11}, antenatal visits during pregnancy and current employment status of mother have significant effect on an indicator of cough (λ_{21}). But, only covariate of antenatal visits during pregnancy has a significant effect on the indicator of diarrhea (λ_{31}) and these results are quite consistent with the previous separate analysis. As further analysis, we excluded the parametric indirect covariates which were nonsignificant in the previous results and include covariates of urban, antenatal visits during pregnancy and current employment status of mother as indirect effects. The results of LVM2 (table 5.26) show that all the parametric indirect covariates have a significant effect on the latent variable "health status" of children.

With regards to non-linear effects on the latent variable health status the patterns are quite similar to the patterns of separate analyses for Nigeria as seen by figure 5.15.

Figure 5.16 displays the results of spatial effects in model LVM1 and LVM2. This suggests that the high risk of all three health status rates is associated with the northeastern part of Nigeria as already indicated by the previous separate results with geoadditive probit models.

5.4.1 Comments

There are both conceptual and technical problems associated with information on prevalence of fever, diarrhea and cough obtained retrospectively from cross-sectional studies. First, seasonal differences of occurrence in diarrhea are difficult to be taken into account in such studies. The researchers think that longitudinal studies may be more appropriate to provide data in different seasons. Second, during the survey, neither the children were examined nor mothers were given a precise definition of what constitutes an episode of various diseases. On the other hand, we have no sufficient information about the children who have died before the survey, and whether the cause of dying was kind of the diseases which are reported here or not. The questions measure (in the DHS) the mother's perception of her child's health rather, than morbidity according to clinical examination. This may create variations among different socio-economic groups because perception of illness is not the same across different social groups. Third, loss of memory of events as well as misinterpretation of the reference period can also contribute to the problems associated with the prevalence of diarrhea (Bateman and Smith, 1991; Gaminiratne, 1991).

Parametrer	Mean	Std	2.5%	97.5%
		Factor Loadings		
1. Fever λ_{11}	1.29*	0.093	1.12	1.46
2. Cough λ_{21}	0.82*	0.04	0.75	0.91
3. Diarrhea λ_{31}	0.79*	0.04	0.71	0.87
		Parametric Indirect Effects		
male	0.135*	0.038	0.059	0.208
anvis	0.218*	0.044	0.131	0.30
trepr	0.088	0.061	-0.034	0.208
work	0.123*	0.05	0.023	0.22
radio	-0.169*	0.05	-0.269	-0.070
educ	-0.061	0.032	-0.125	0.001
toilet	-0.129	0.094	-0.319	0.051
		Semi-Parametric Indirect Effects		
Chage	0.059*	0.043	0.014	0.169
BMI	0.017*	0.028	0.000	0.085
Mageb	0.004*	0.011	0.0003	0.019
reg	0.201*	0.112	0.063	0.484
		Parametric Direct Effects		
urban(a_{11})	0.0329	0.07	-0.1	0.17
elect(a_{12})	-0.36	0.274	-0.89	0.182
water(a_{13})	0.129	0.089	-0.041	0.30
urban(a_{21})	0.152*	0.054	0.044	0.25
elect(a_{22})	-0.071	0.23	-0.52	0.39
water(a_{23})	-0.016	0.071	-0.16	0.124
urban(a_{31})	-0.207*	0.056	-0.317	-0.09
elect(a_{32})	0.012	0.22	-0.42	0.47
water(a_{33})	0.042	0.072	-0.098	0.187

Table 5.22: Results of LVM1 including direct and indirect effects for Egypt.
(*: Statistically significant at 2.5%)

Parametrer	Mean	Std	2.5%	97.5%
		Factor Loadings		
1. Fever λ_{11}	1.285*	0.087	1.129	1.455
2. Cough λ_{21}	0.827*	0.043	0.748	0.921
3. Diarrhea λ_{31}	0.789*	0.043	0.703	0.874
		Parametric Indirect Effects		
male	0.136*	0.037	0.063	0.208
anvis	0.219*	0.043	0.136	0.306
trepr	0.090	0.063	-0.034	0.210
work	0.124*	0.049	0.027	0.229
radio	-0.17^*	0.050	-0.269	-0.007
educ	-0.062	0.032	-0.127	0.006
toilet	-0.132	0.094	-0.319	0.049
elect	-0.152	0.187	-0.528	0.207
water	0.058	0.056	-0.055	0.169
		Semi-Parametric Indirect Effects		
Chage	0.060*	0.0433	0.015	0.173
BMI	0.015*	0.022	0.0008	0.076
Mageb	0.002*	0.0034	0.0003	0.011
reg	0.199*	0.102	0.072	0.457
		Parametric Direct Effects		
urban(a_{11})	0.040	0.064	-0.087	0.165
urban(a_{21})	0.138*	0.052	0.031	0.240
urban(a_{31})	-0.207^*	0.054	-0.315	-0.095

Table 5.23: Results of LVM2 including direct and indirect effects for Egypt

Parametrer	Mean	Std	2.5%	97.5%
		Factor Loadings		
1. Fever λ_{11}	1.273*	0.099	1.090	1.487
2. Cough λ_{21}	0.824*	0.040	0.746	0.911
3. Diarrhea λ_{31}	0.796*	0.047	0.706	0.889
		Parametric Indirect Effects		
male	0.135*	0.038	0.060	0.209
anvis	0.22*	0.044	0.138	0.313
work	0.127*	0.050	0.02	0.225
radio	-0.186^*	0.049	-0.286	-0.090
educ	-0.065	0.033	-0.129	0.008
		Semi-Parametric Indirect Effects		
Chage	0.0597*	0.043	0.014	0.175
BMI	0.016*	0.027	0.0008	0.088
Mageb	0.003*	0.0056	0.0003	0.001
reg	0.202*	0.106	0.069	0.0473
		Parametric Direct Effects		
urban(a_{11})	0.041	0.066	-0.088	0.171
urban(a_{21})	0.141*	0.055	0.0318	0.248
urban(a_{31})	-0.209^*	0.056	-0.316	-0.09

Table 5.24: Results of LVM3 including direct and indirect effects for Egypt

Parametrer	Mean	Std	2.5%	97.5%
Factor Loadings				
1. Fever λ_{11}	0.998*	0.084	0.845	1.162
2. Cough λ_{21}	0.917*	0.072	0.77	1.063
3. Diarrhea λ_{31}	0.753*	0.056	0.647	0.862
Parametric Indirect Effects				
male	0.049	0.068	-0.083	0.185
educ	0.0008	0.054	-0.104	0.110
toilet	-0.081	0.071	-0.219	0.058
radio	-0.028	0.040	-0.103	0.052
trepr	0.222*	0.076	0.070	0.370
Semi-Parametric Indirect Effects				
Chage	0.051*	0.050	0.0107	0.181
BMI	0.005*	0.012	0.0003	0.032
Mageb	0.003*	0.0048	0.0003	0.017
reg	0.437*	0.156	0.211	0.812
Parametric Direct Effects				
urban(a_{11})	-0.224*	0.084	-0.387	-0.059
anvis(a_{12})	0.014	0.084	-0.150	0.180
elect(a_{13})	0.013	0.084	-0.150	0.177
work(a_{14})	0.030	0.075	-0.114	0.182
water(a_{15})	0.048	0.050	-0.052	0.146
urban(a_{21})	-0.114	0.083	-0.280	0.048
anvis(a_{22})	0.255*	0.083	0.092	0.416
elect(a_{23})	0.046	0.081	-0.115	0.209
work(a_{24})	0.161*	0.070	0.023	0.300
water(a_{25})	-0.032	0.049	-0.130	0.065
urban(a_{31})	-0.038	0.080	-0.199	0.119
anvis(a_{32})	-0.274*	0.076	-0.420	-0.122
elect(a_{33})	-0.032	0.079	-0.184	0.124
work(a_{34})	-0.033	0.067	-0.161	0.098
water(a_{35})	-0.037	0.049	-0.135	0.058

Table 5.25: Results of LVM1 including direct and indirect effects for Nigeria

Parametrer	Mean	Std	2.5%	97.5%
Factor Loadings				
1.Fever λ_{11}	1.014*	0.078	0.871	1.189
2.Cough λ_{21}	0.945*	0.076	0.808	1.112
3.Diarrhea λ_{31}	0.730*	0.056	0.621	0.842
Parametric Indirect Effects				
trepr	0.217*	0.071	0.010	0.1623
work	0.054*	0.077	0.0004	0.024
urban	-0.154*	0.072	0.0003	0.014
anvis	0.0026*	0.086	0.205	0.770
Semi-Parametric Indirect Effects				
Chage	0.048*	0.0438	0.010	0.162
BMI	0.0049*	0.007	0.0004	0.0244
Mageb	0.0028*	0.004	0.0003	0.0145
reg	0.421*	0.144	0.205	0.770
Parametric Direct Effects				
elect(a_{11})	0.007	0.087	-0.168	0.174
water(a_{12})	0.053	0.051	-0.045	0.155
male(a_{13})	0.035	0.067	-0.091	0.167
educ(a_{14})	0.022	0.058	-0.093	0.138
toilet(a_{15})	-0.194*	0.066	-0.327	-0.065
radio(a_{16})	-0.0140	0.041	-0.093	0.071
elect(a_{21})	0.0371	0.082	-0.120	0.201
water(a_{22})	-0.035	0.050	-0.131	0.0654
male(a_{23})	0.0013	0.066	-0.125	0.133
educ(a_{24})	0.045	0.055	-0.063	0.154
toilet(a_{25})	0.081	0.064	-0.040	0.210
radio(a_{26})	0.009	0.040	-0.068	0.0915
elect(a_{31})	-0.0124*	0.077	-0.164	-0.062
water(a_{32})	-0.024	0.047	-0.115	0.066
male(a_{33})	0.114	0.062	-0.004	0.235
educ(a_{34})	-0.098	0.055	-0.206	0.0095
toilet(a_{35})	-0.120	0.063	-0.248	0.0024
radio(a_{36})	-0.075	0.037	-0.148	0.0045

Table 5.26: Results of LVM2 including direct and indirect effects for Nigeria

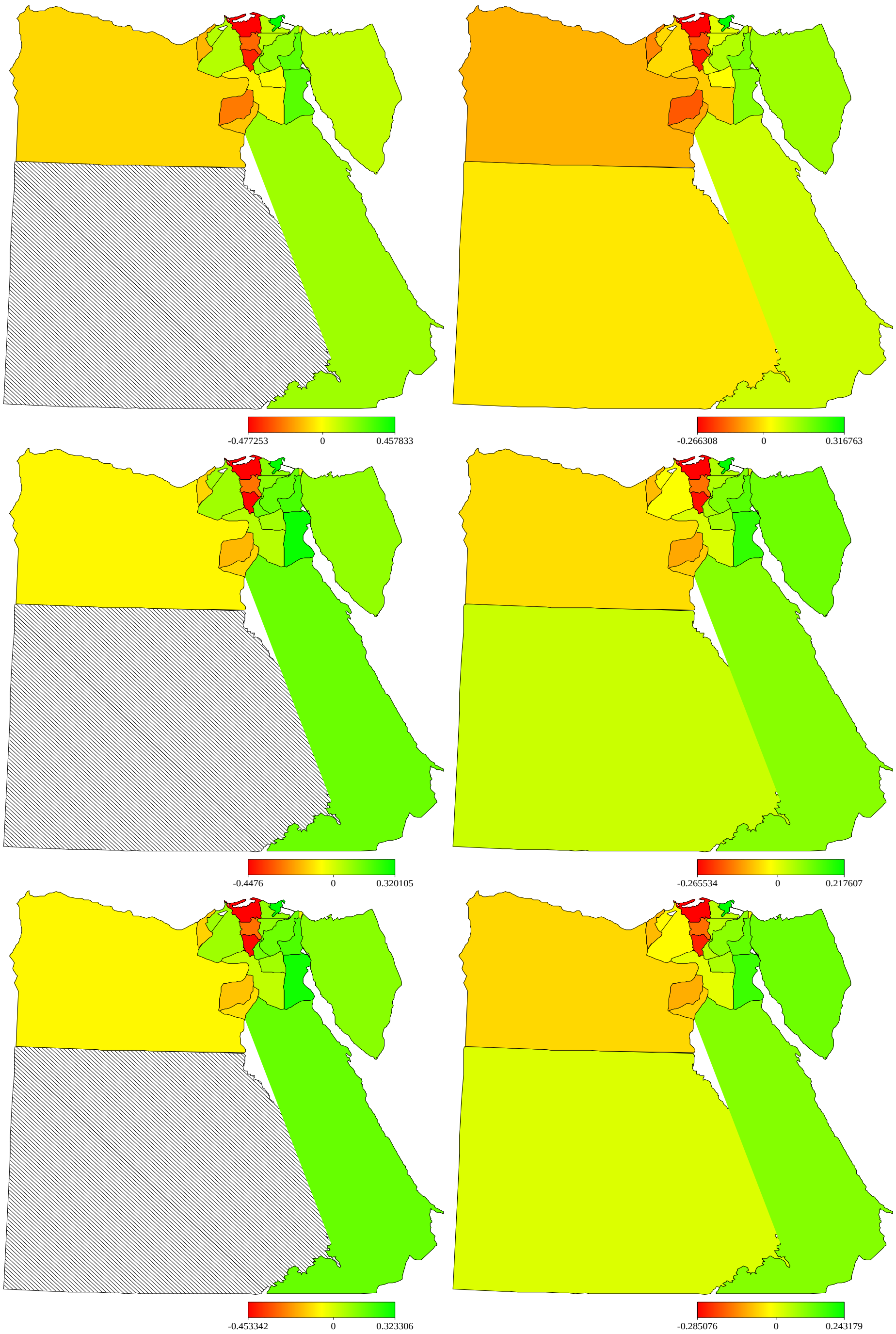


Figure 5.1: Maps of Egypt for diarrhea showing unstructured (top left) and structured (right left) spatial effects (for model M0), unstructured (second from top left) and structured (second from top right) spatial effects (for model M1), unstructured (bottom left) and structured (bottom right) spatial effects (for model M2) using probit model .

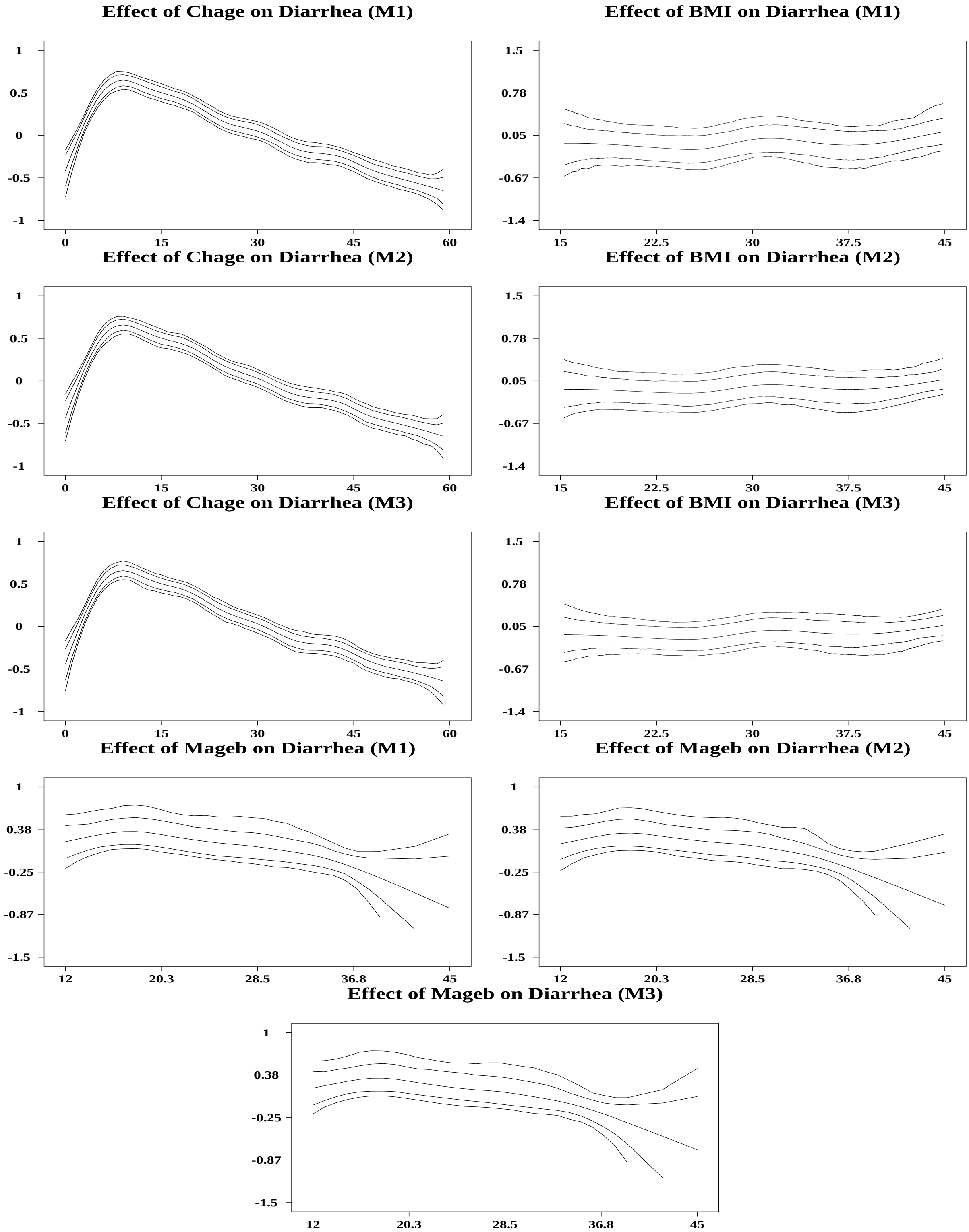


Figure 5.2: Non-linear effects from top to bottom: child's age, mother's BMI (for model M1), child's age and mother's BMI (for model M2), child's age and mother's BMI (for model M3) and mother's age (for M1, M2 and M3) on diarrhea for Egypt using probit model.

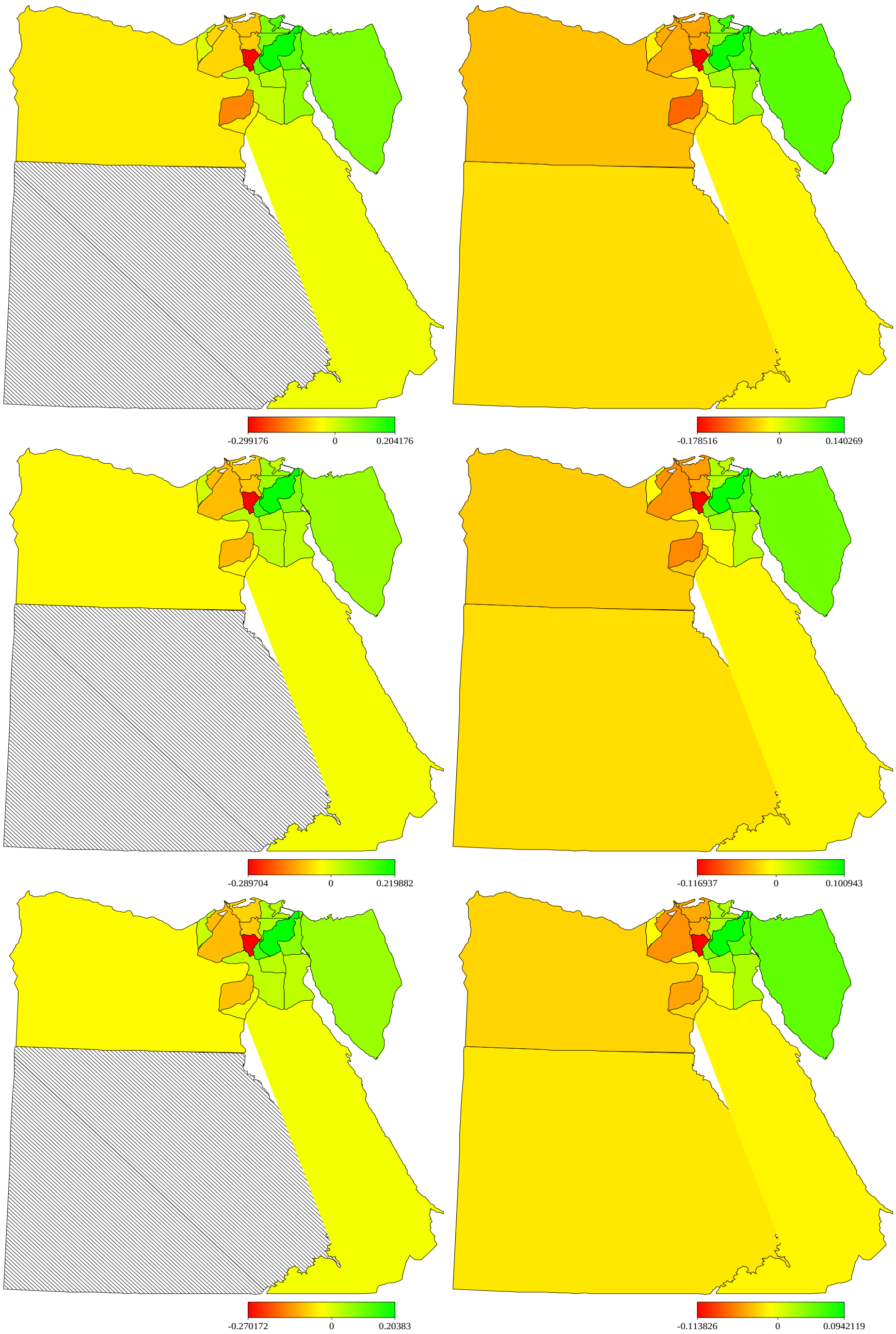


Figure 5.3: Maps of Egypt for fever showing unstructured (top left) and structured (right left) spatial effects (for model M0), unstructured (second from top left) and structured (second from top right) spatial effects (for model M1), unstructured (bottom left) and structured (bottom right) spatial effects (for model M2) using probit model.

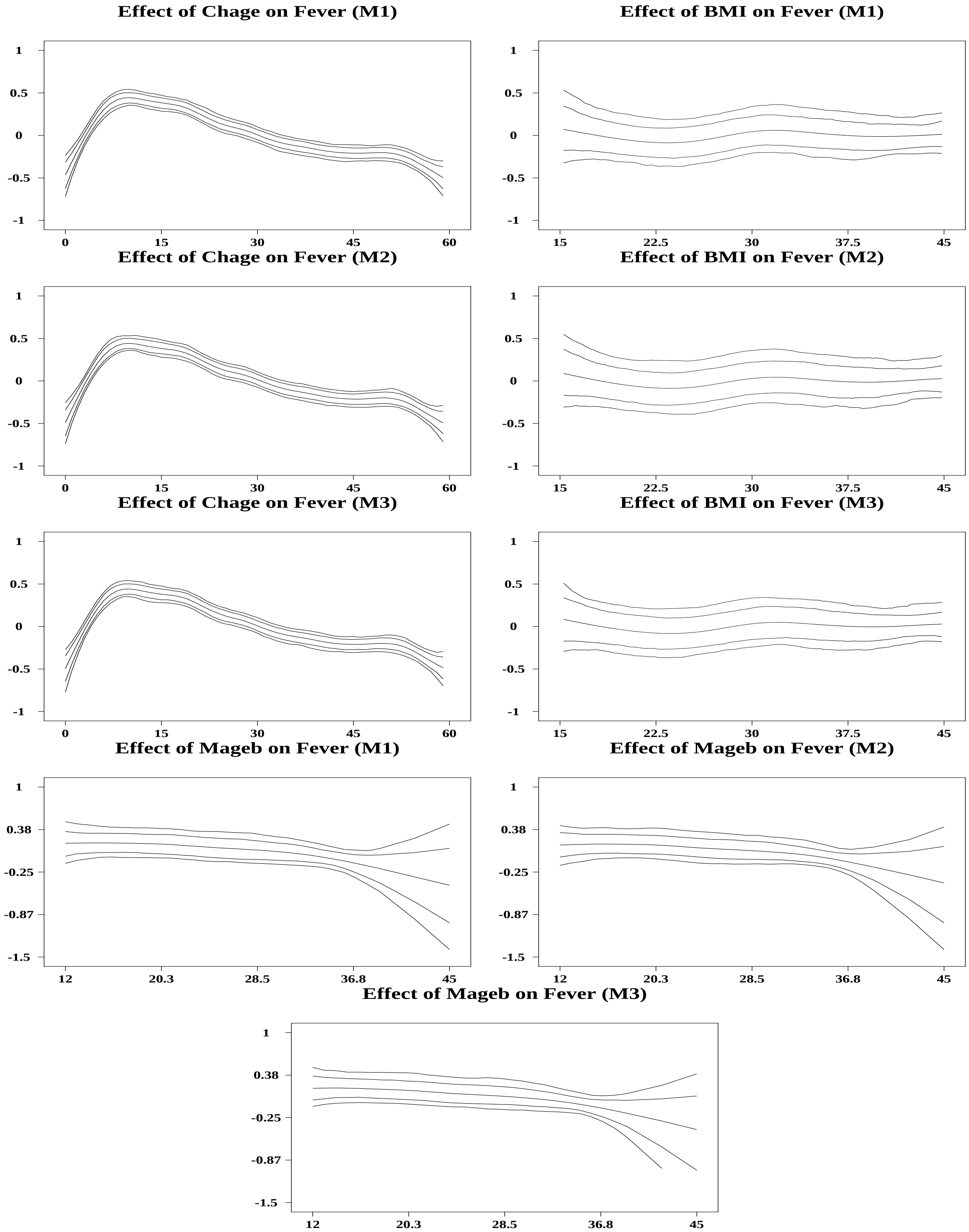


Figure 5.4: Non-linear effects from top to bottom: child's age and mother's BMI (for model M1), child's age and mother's BMI (for model M2), child's age and mother's BMI (for model M3) and mother's age (for M1, M2 and M3) on fever for Egypt using probit model.

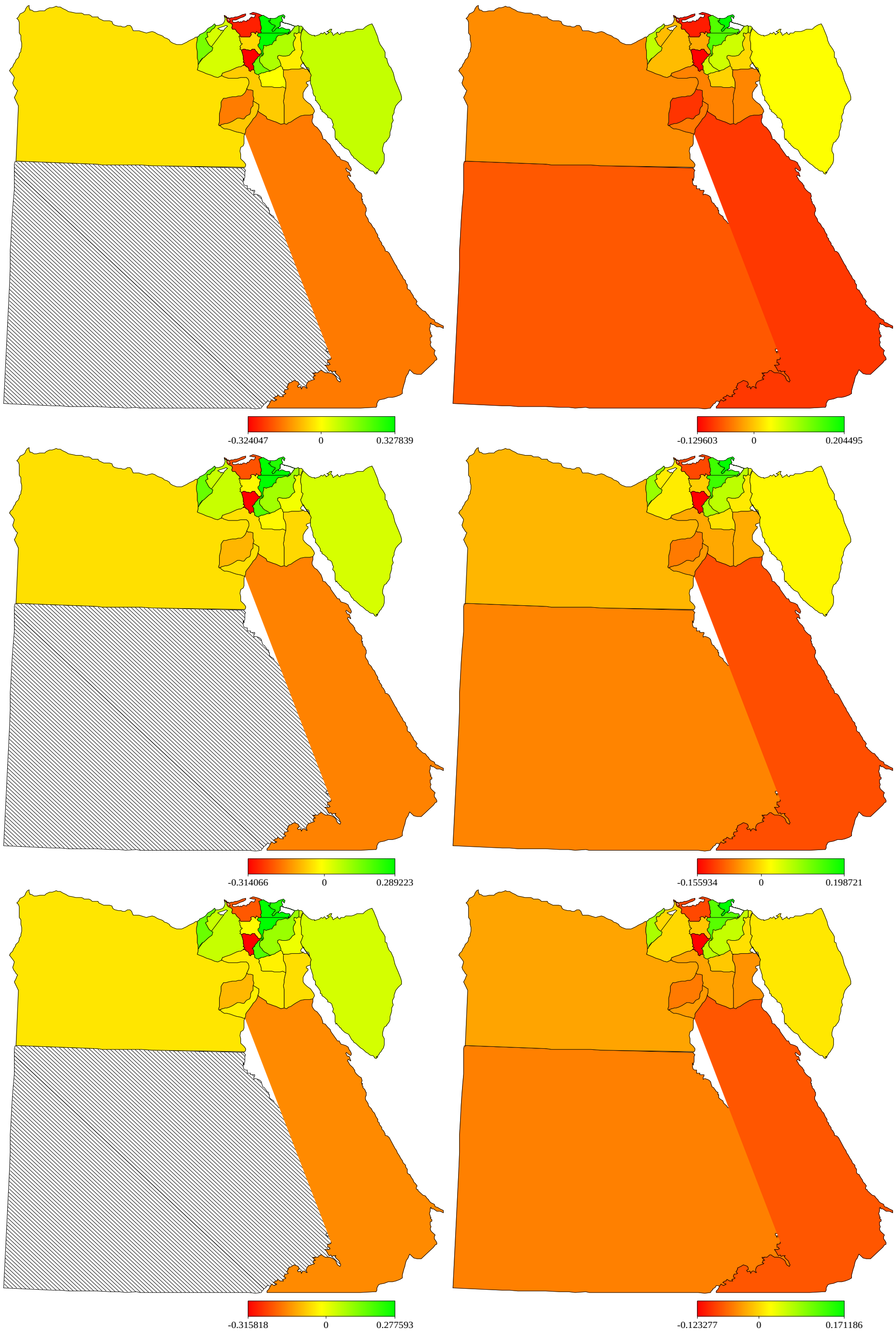


Figure 5.5: Maps of Egypt for cough showing unstructured (top left) and structured (right left) spatial effects (for model M0), unstructured (second from top left) and structured (second from top right) spatial effects (for model M1), unstructured (bottom left) and structured (bottom right) spatial effects (for model M2) using probit model .

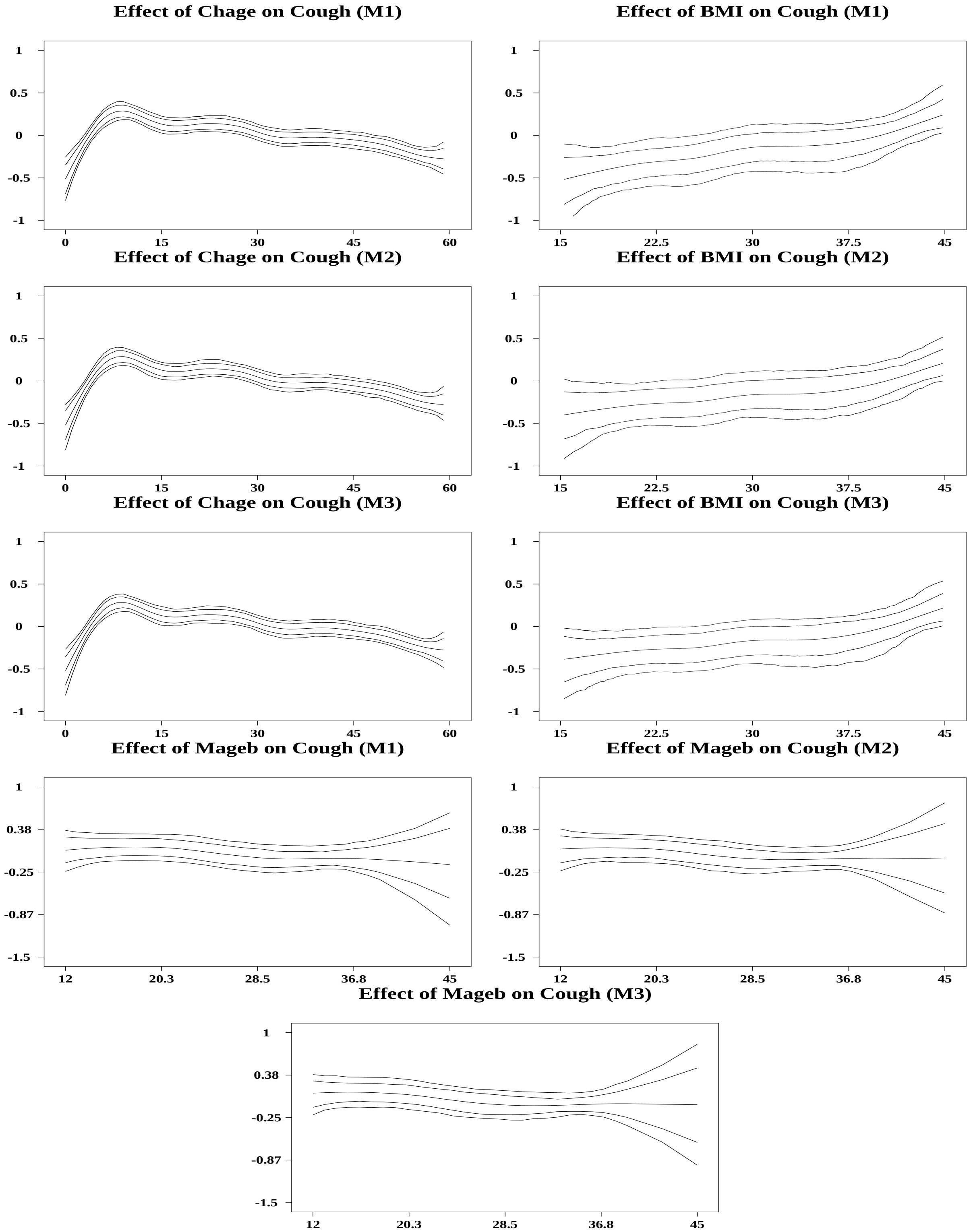


Figure 5.6: Non-linear effects from top to bottom: child's age and mother's BMI (for model M1), child's age and mother's BMI (for model M2), child's age and mother's BMI (for model M3) and mother's age (for M1, M2, and M3) on cough for Egypt using probit model.

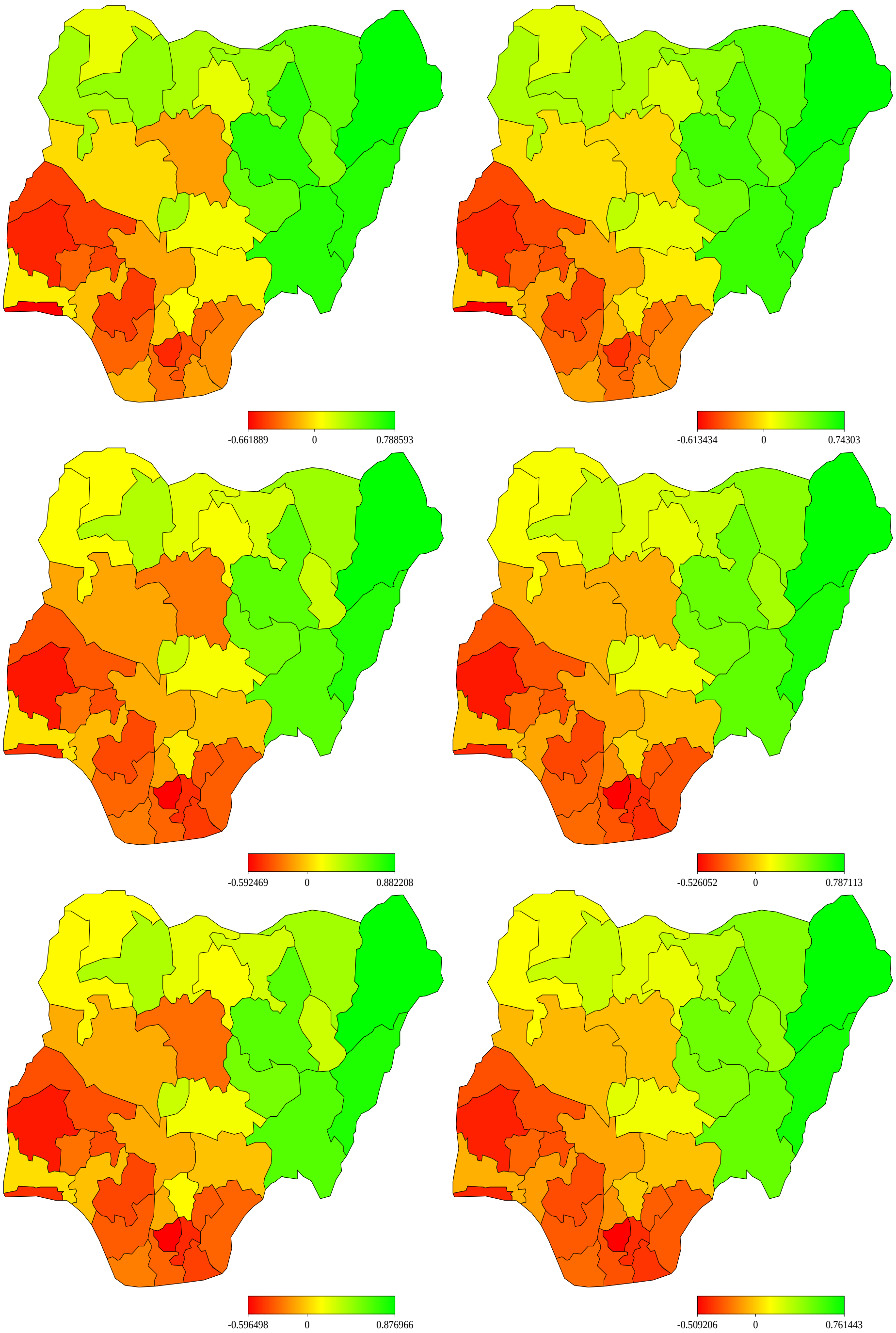


Figure 5.7: Maps of Nigeria for diarrhea showing unstructured (top left) and structured (right left) spatial effects (for model M0), unstructured (second from top left) and structured (second from top right) spatial effects (for model M1), unstructured (bottom left) and structured (bottom right) spatial effects (for model M2) using probit model .

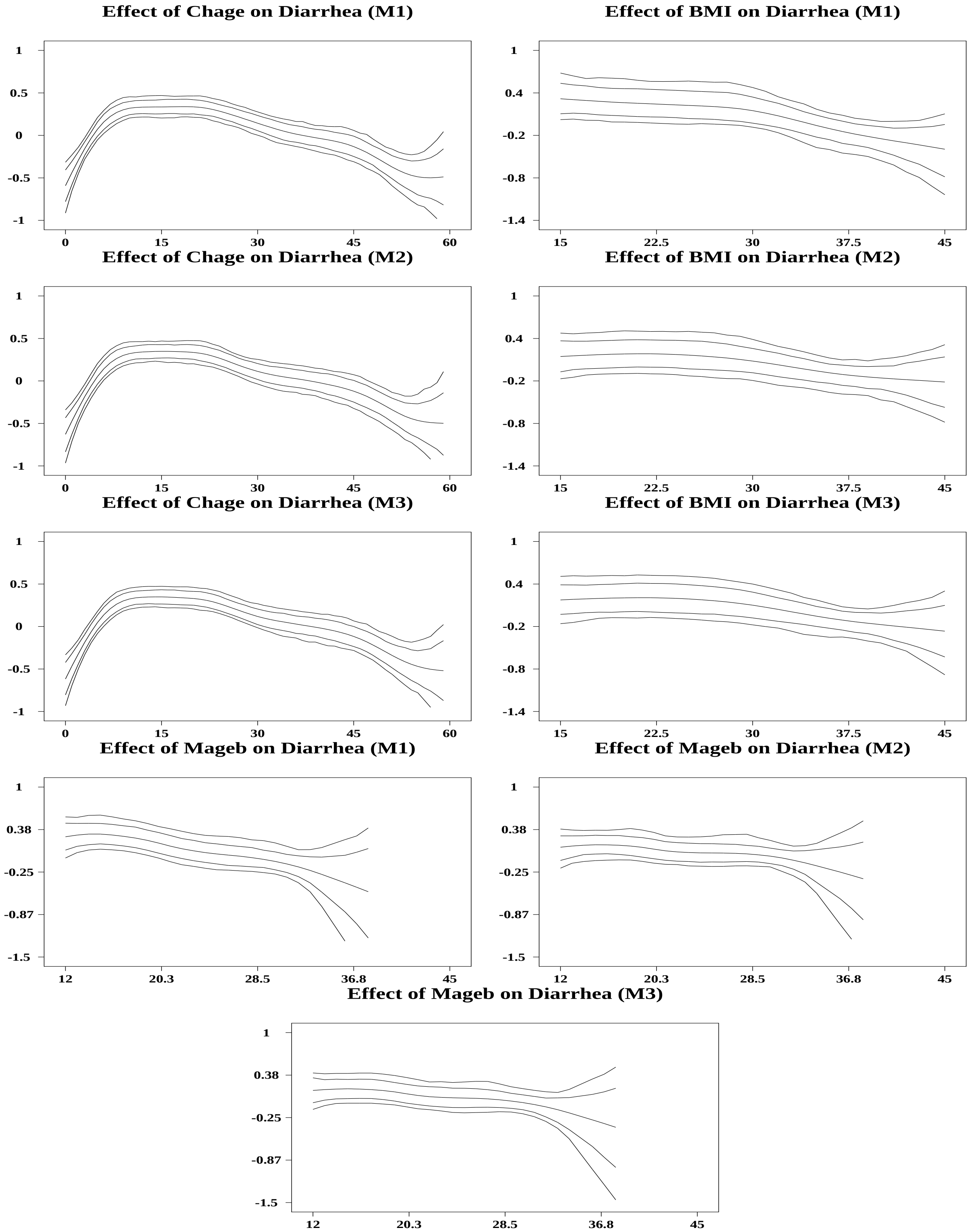


Figure 5.8: Non-linear effects from top to bottom: child's age and mother's BMI (for model M1), child's age and mother's BMI (for model M2), child's age and mother's BMI (for model M3) and mother's age (for M1, M2, and M3) on diarrhea for Nigeria using probit model.

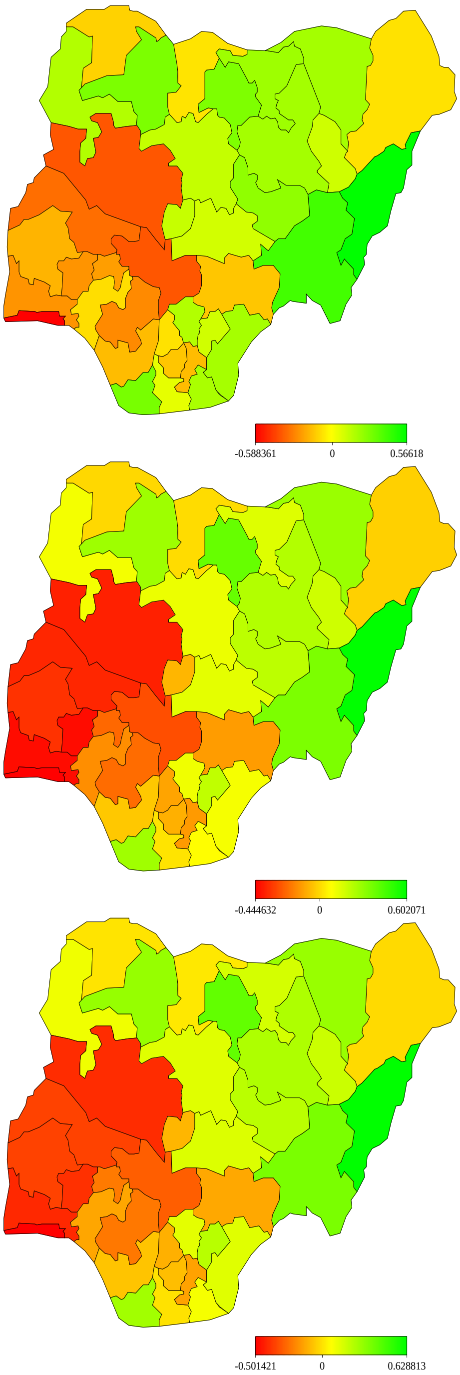


Figure 5.9: Maps of Nigeria for fever showing unstructured (top left) and structured (right left) spatial effects (for model M0), unstructured (second from top left) and structured (second from top right) spatial effects (for model M1), unstructured (bottom left) and structured (bottom right) spatial effects (for model M2) using probit model.

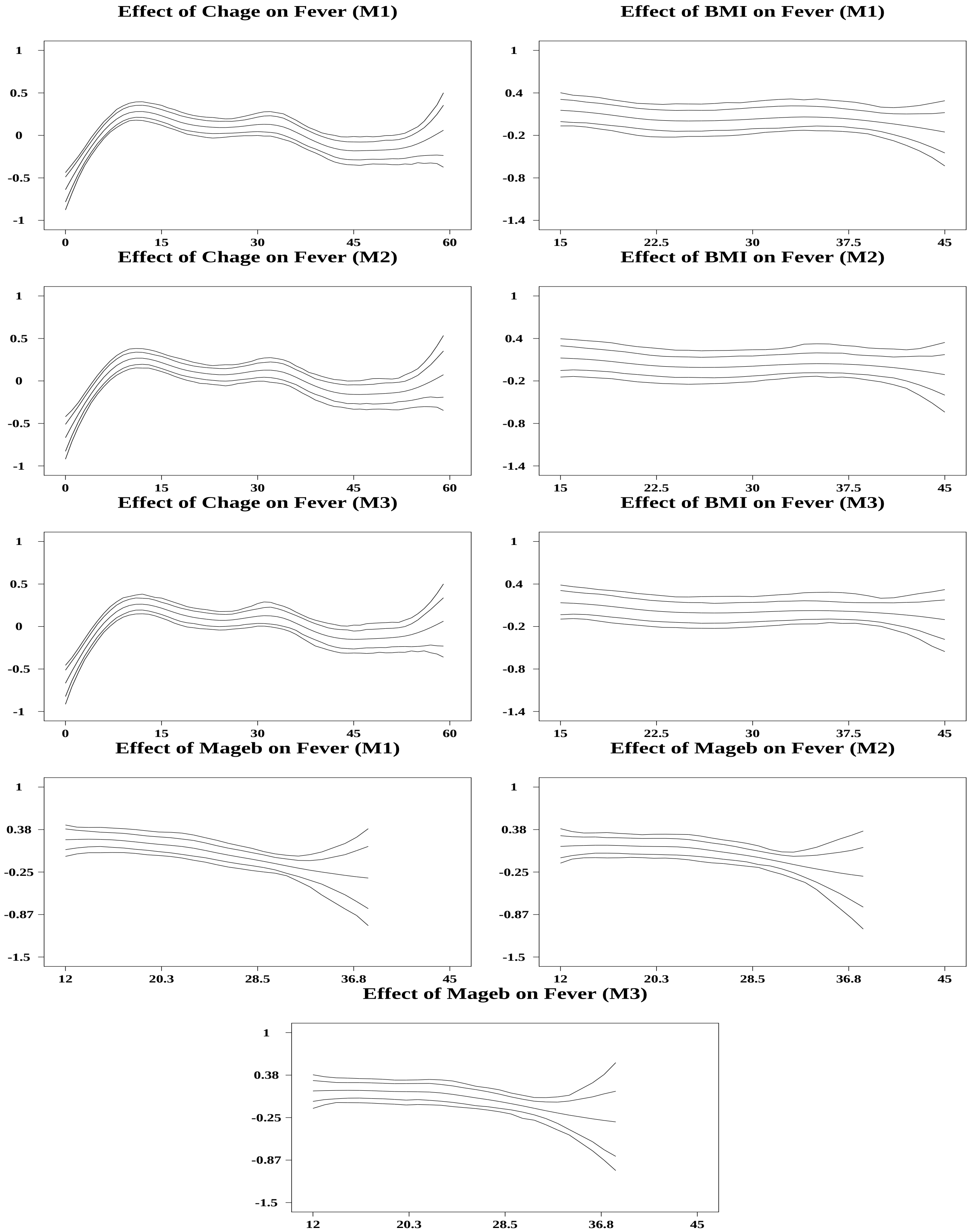


Figure 5.10: Non-linear effects from top to bottom: child's age and mother's BMI (for model M1), child's age and mother's BMI (for model M2), child's age and mother's BMI (for model M3) and mother's age (for M1, M2, and M3) on fever for Nigeria using probit model.

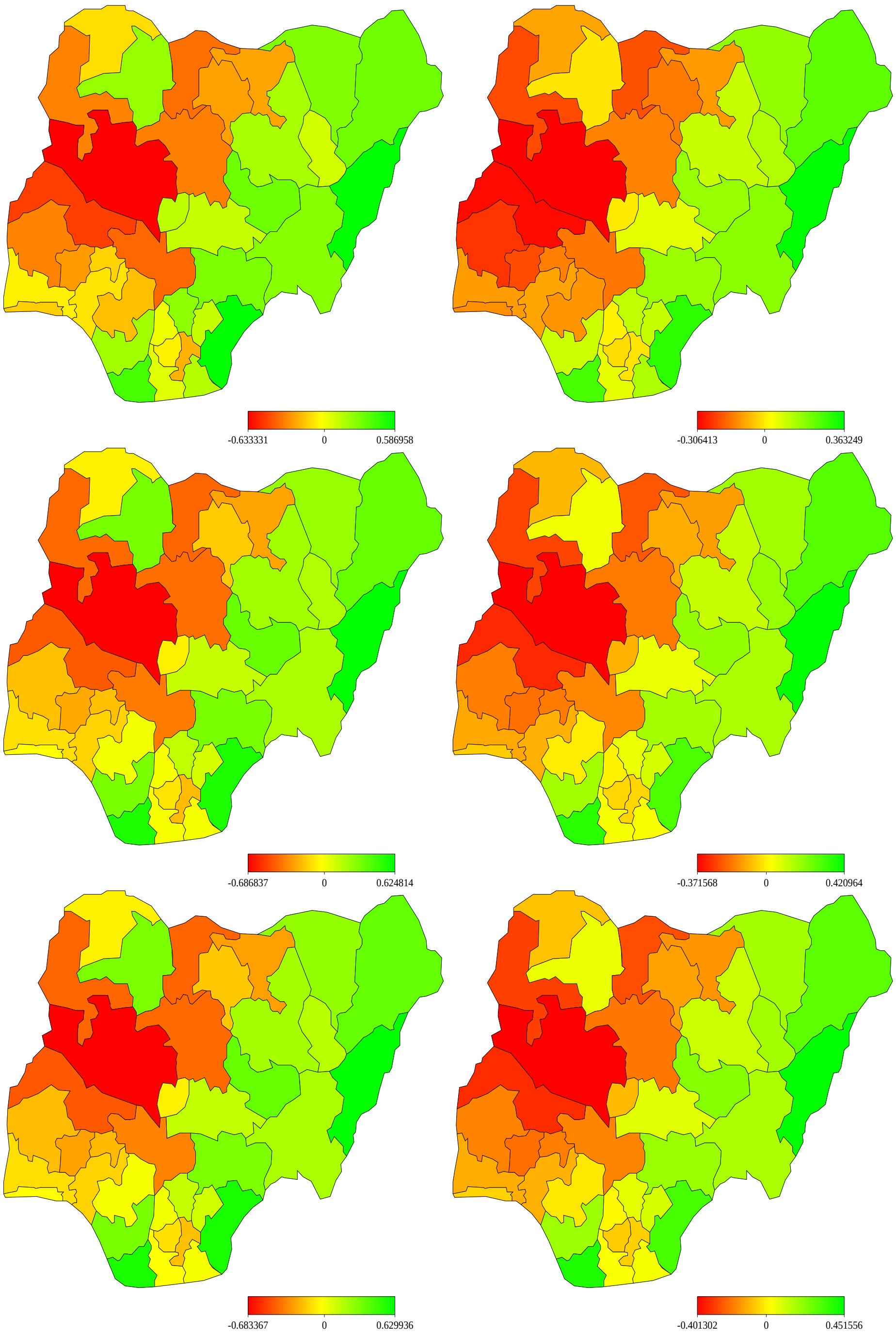


Figure 5.11: Maps of Nigeria for cough showing unstructured (top left) and structured (right left) spatial effects in model M0, unstructured (second from top left) and structured (second from top right) spatial effects in model M1, unstructured (bottom left) and structured (bottom right) spatial effects in model M2 using probit model.

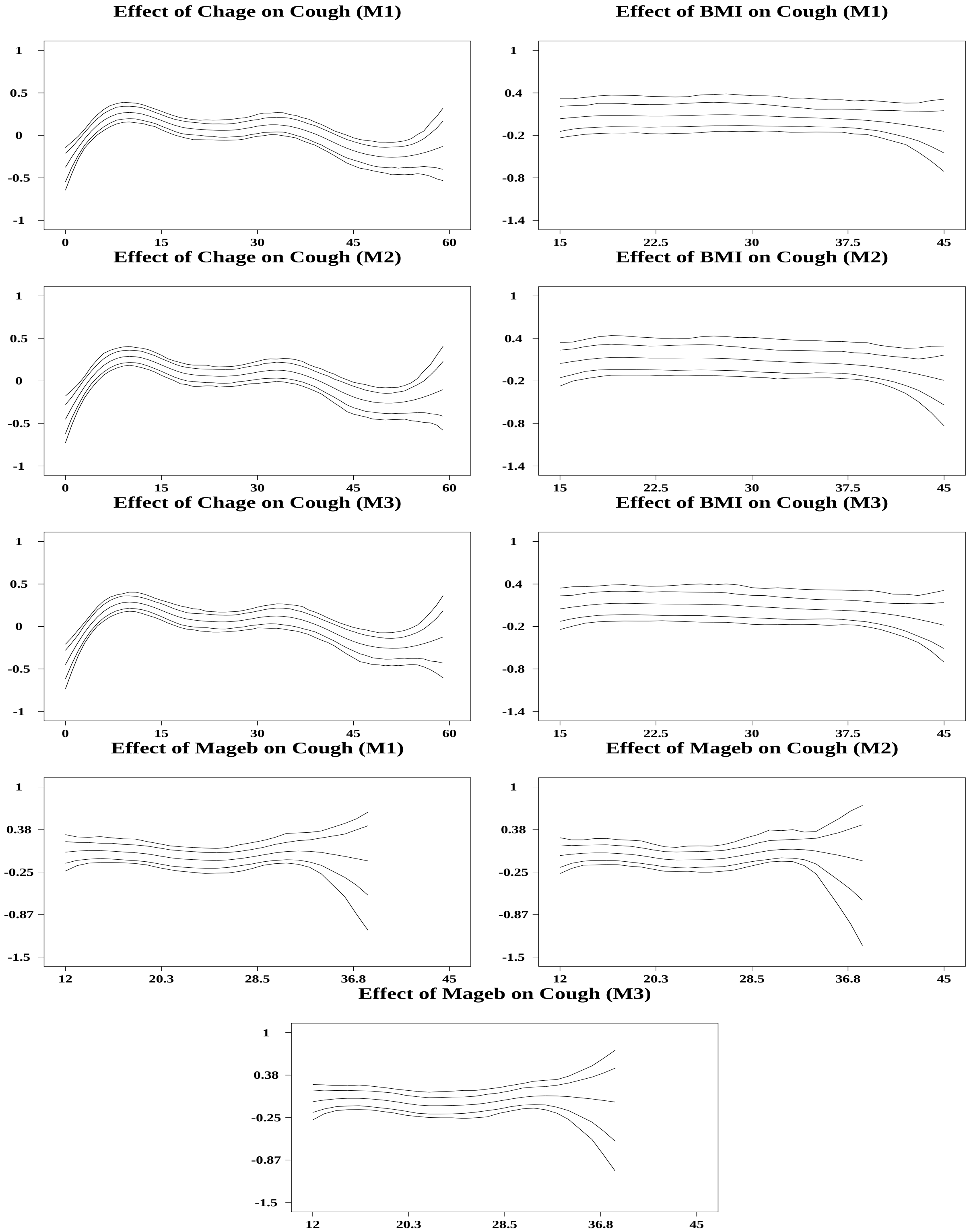


Figure 5.12: child's age and mother's BMI (for model M1), child's age and mother's BMI (for model M2), child's age and mother's BMI (for model M3) and mother's age (for M1, M2, and M3) on cough for Nigeria using probit model.

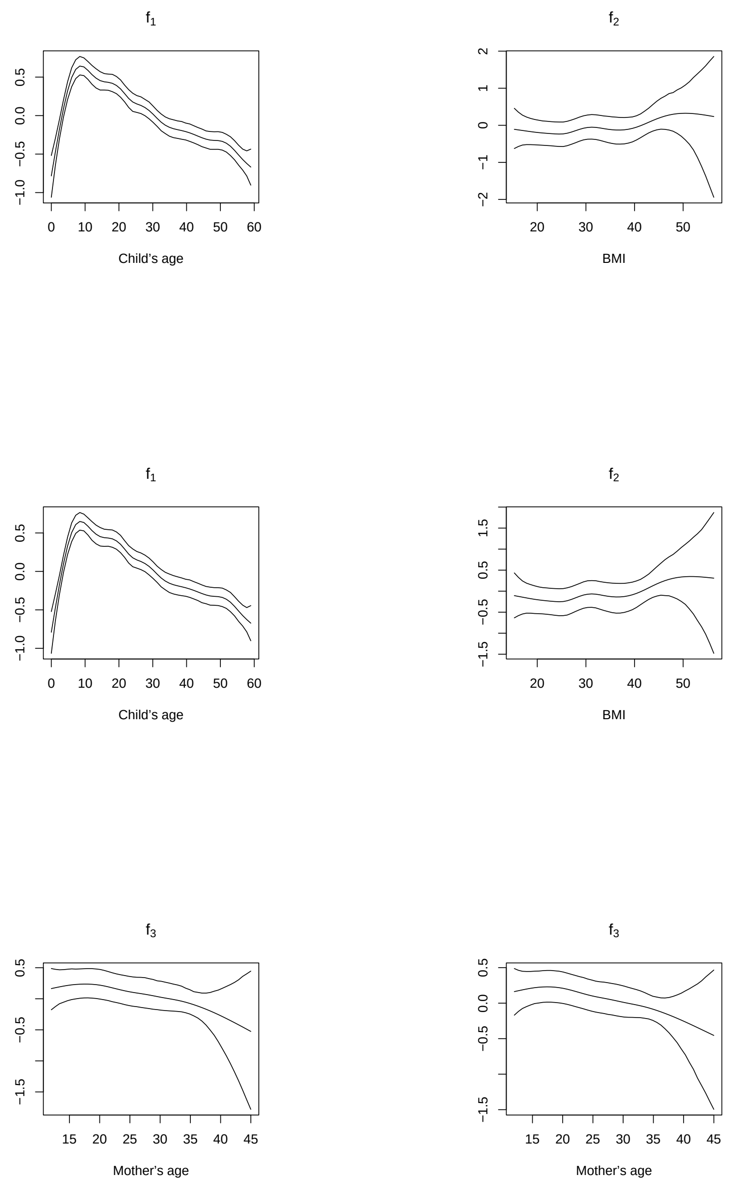


Figure 5.13: Non-linear effects from top to bottom: child's age , mother's BMI and mother's age at birth (for model LVM1), child's age, mother's BMI and mother's age at birth (for model LVM2) on the indicators of a latent variable "health status" of children disease for Egypt using Bayesian latent variable model for binary responses.

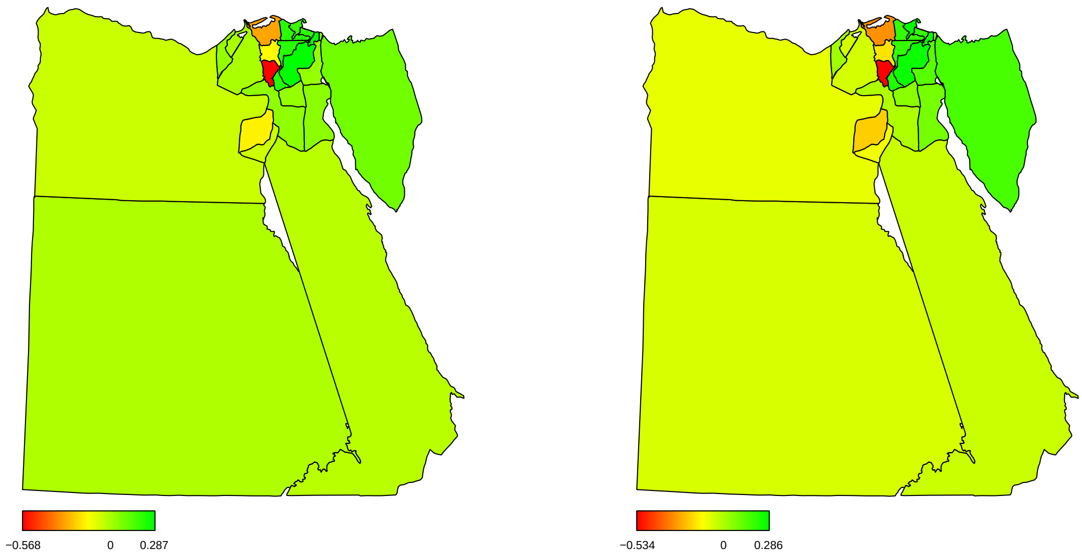


Figure 5.14: Posterior mean for latent variable model for LVM1 (left panel) and LVM2 (right panel) on diseases in Egypt.

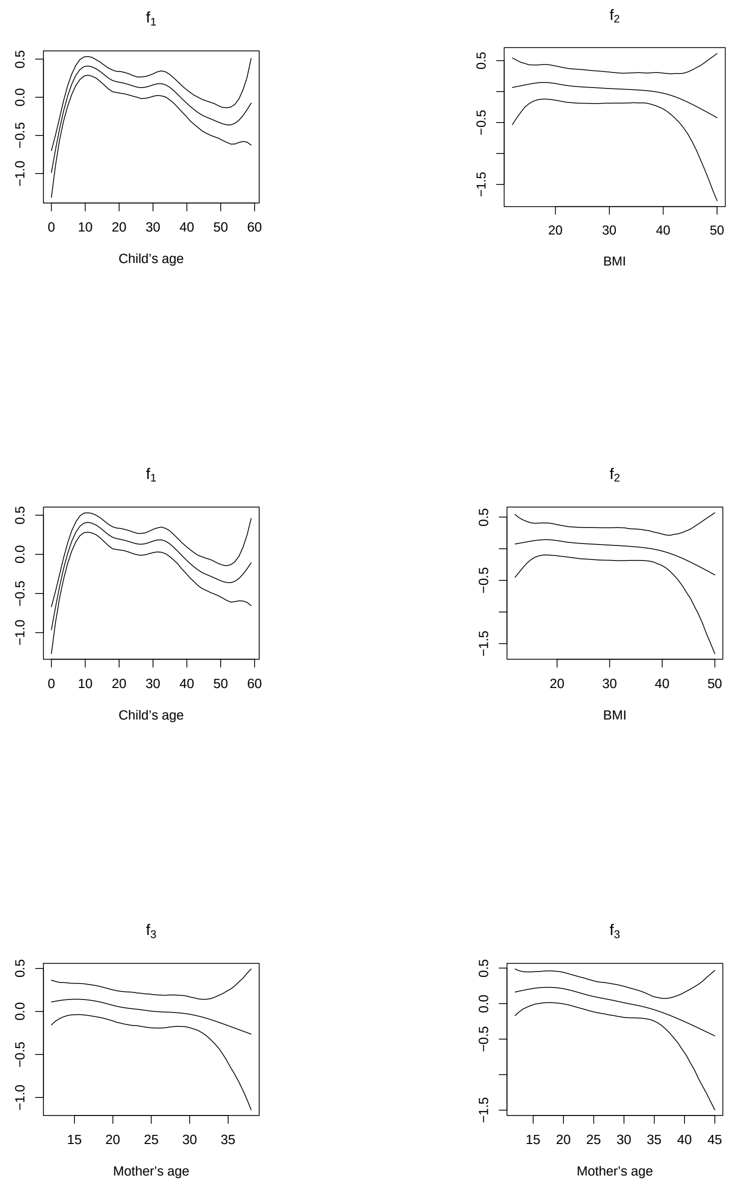


Figure 5.15: Non-linear effects from top to bottom: child's age , mother's BMI and mother's age at birth (for model LVM1), child's age , mother's BMI and mother's age at birth(for model LVM2) on the indicators of a latent variable "health status" of children disease for Nigeria using Bayesian latent variable model for binary responses.

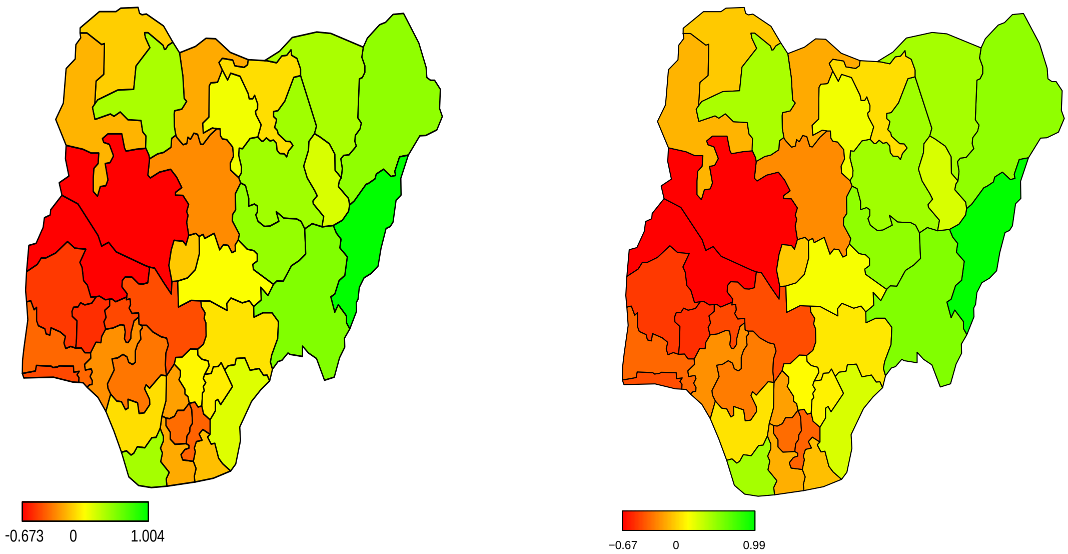


Figure 5.16: Posterior mean for latent variable model for LVM1 (left panel) and LVM2 (right panel) on diseases in Nigeria.

Chapter 6

Semiparametric Modelling of Malnutrition Status of Children using Geoadditive Gaussian Regression and Latent Variable Models

Abstract

In this chapter, we investigate the geographical and socioeconomic determinants of childhood undernutrition in Egypt and Nigeria using the 2003 DHS. We use geoadditive Gaussian regression and latent variable models to explore models of the effects of selected socioeconomic covariates. In the first step, we use separate geoadditive Gaussian models with the continuous responses variables stunting (*height-for-age*), underweight (*weight-for-age*) and wasting (*weight-for-height*) as indicators of nutritional status in our case study. In a second step, based on the results of the first step, we apply the geoadditive Gaussian latent variable model for continuous indicators in which the three measurements of the malnutrition status of children are assumed as indicators for the latent variable "nutritional status".

6.1 Introduction

The previous studies have distinguished between immediate, intermediate, and underlying determinants (see UNICEF, 2003). Whilst undernutrition is always immediately related to either insufficient nutrient intake or the inability of the body to absorb nutrients (primarily due to illness), these are themselves caused food security, care practices, and the health environment at the household level, which are influenced by the socioeconomic and demographic situation of households and communities (UNICEF, 2003; Smith and Haddad, 1999; Klasen, 1999). In order to capture this complex chain of causation, previous studies have focused on a particular level of causality (e.g. Smith and Haddad, 1999; Moradi, 1999; Pelletier, 1999). The others have either estimated structural equations that address the interactions (e.g. Guilkey and Riphahn, 1998), have also used graphical chain models to assess the causal pathways (Caputo et al., 2002), or multi-level modelling techniques (e.g. Nyovani et al., 1999). We will use mainly models factors which underlie determinants of undernutrition (see Kandala, Fahrmeir and Klasen, 2002). The included covariates are measures of the following: household resources (access to electricity and radio), access to water and sanitation, mother's education and employment status, mother's age, BMI, the child's age, sex and the place of residence (urban or rural).

In this chapter, we investigate the geographical and socioeconomic determinants of childhood undernutrition in Egypt and Nigeria. We investigate the impact of the various bio-demographic and socio-economic variables on the indicators of undernutrition using flexible semiparametric models as done in the analysis of childhood diseases. To build a regression model for undernutrition, we first have to define a distribution for the response variable. In this application, it is reasonable to assume that Z-score is Gaussian distributed; thus in principle, model 6.2 could be applied. The analysis started by employing a separate geoadditive Gaussian model to continuous response variables for wasting, stunting and underweight. We then apply geoadditive latent variable models, based on these results, where the three undernutrition variables are taken as indicators for the latent variable malnutrition of a child. This chapter consists of four sections. Section 1 describes the geoadditive Gaussian and geoadditive latent model, whilst section 2 contains statistical inference and results using the geoadditive and the latent variable models. Section 3 includes discussion and comments.

6.2 Bayesian Geoadditive Regression and Latent Variable Models of Childhood Malnutrition

In the following, we focus on geoadditive Gaussian model for continuous response variables to analyze the effects of metrical, categorical, and spatial covariates on the stunting, wasting and underweight response variable in the separate analysis. Furthermore, we use "nutritional status" as the indicator in the analysis of the latent variable models.

6.2.1 Geoadditive Gaussian Regression

In this section, we concentrate on separate analyses for three types of anthropometric status of the child, with most of the research focused on children below five years of age in both countries, using a flexible regression method to model the effect of covariates that have linear and nonlinear effects, and the effect of geographical covariate on the three types of undernutrition (stunting, wasting, and underweight). In our application, the responses of childhood undernutrition are stunting, wasting and underweight, which are measured as standardized Z-scores. Traditionally, the effect of the covariates on the response is modelled by a linear predictor:

$$\eta_{ij}^{lin} = x'_{ij}\beta_j + w'_{ij}\gamma_j \quad j = 1, .., 3, \tag{6.1}$$

where observations (x_i, w_i), i=1,..,n, on a metrical response y, a vector $x = (x_1, .., x_p)$ of metrical covariates and vector $w = (w_1, ..w_k)$ of categorical covariates.

In our analysis, nonlinear effects of the spatial structure can be included, using regional dummy variables (see below). In this work, particular emphasis is on the nonlinear effects of the metrical covariates (age of the child) Chage, BMI (mother's body mass index) and Mageb (age of mother at birth), categorical covariates (male, urban, having radio, etc.) and the spatial covariate (child's district of residence) on childhood undernutrition. Thus, we replace the strictly linear predictor 6.1 by the more flexible geoadditive predictor

$$\eta_{ij} = \beta_{0j} + f_1(Chage_i) + f_2(BMI_i) + f_3(Mageb_i) + f_{spat_i}(s) + \omega_i^{\prime}\gamma_j \quad (6.2)$$

where w includes the categorical covariates in effect coding. The function f_1, f_2 and f_3 are non-linear smooth effects of the metrical covariates which are modelled by P-splines priors and f_{spat} is the effect of the spatial covariate $s_i \in 1; ...; S$ labeling the districts in both countries. Regression models with predictors as in 6.2 are referred to as geoadditive models. In a further step and as usual we split up the spatial effect f_{spat} into a spatially correlated (structured) effect modelled by a Markov field prior (Beseg, York and Mallie, 1990) and uncorrelated (unstructured) effects which are assumed to be i.i.d. random variables, as done in the analysis of childhood diseases. In a Bayesian approach unknown functions f_j and γ_j as well as the variance parameter σ^2 are considered as random variables and have to be supplemented with appropriate prior assumptions as developed by Fahrmeir and Lang (2001) and Brezeger and Lang (2005); see also chapter 2 of the current work.

For further analysis and to capture the drawback of separate Gaussian models for each of the continuous responses, we will use the latent variables models as seen later.

6.2.2 Latent Variable Model for Continuous Responses

In this application, we introduce latent variable v, which reflects undernutrition status; and as done in the analysis of childhood disease, we assume only a one-dimensional latent variable in the preliminary analysis. The LVM is extend later on to multi-dimensional latent variables with different types of observable responses (see Raach, 2005 and Fahrmeir and Raach, 2006).

The latent variable model (LVM) for Gaussian responses $y_j, j = 1, .., p$ and scalar υ is given through the *Gaussian measurement model*:

$$y_{ij} = \lambda_0 + a_j' w_i + \lambda_j \upsilon_i + \epsilon_{ij}, i = 1, .., n \, , j = 1, .., p, \tag{6.3}$$

with i.i.d, Gaussian errors ϵ_{ij}. In this model, υ_i is the unobservable value of individual i, λ_j is the "factor loading", and $\lambda_j \upsilon_i$ is the effect of υ_i. In addition, w_i are the direct effects which affect the observed variables directly and a_j is the matrix of regression coefficients. The restriction to σ_υ=var$(\upsilon) = 1$ is necessary as discussed before (chapter 5) for identifiable reasons. Continuous variables are observed directly, hence the underlying variable is obsolete.

The general form of geoadditive structural model for continuous response is:

$$\upsilon_i = u_i' \alpha + f_1(x_{i1}) + ... + f(x_{iq}) + f_{geo}(s_i) + \delta_i, \tag{6.4}$$

with i.i.d. Gaussian errors $\delta_i \sim N(0, 1)$. In this application, attention is restricted to Gaussian LVMs with continuous response variables.

6.3 Statistical Inference and Results

We present a unified approach for Bayesian inference for each type of undernutrition; stunting, wasting and underweight are analyzed separately via Markov chain Monte Carlo in geoadditive models as the first step of our analysis. Different types of covariates, such as the usual covariates with fixed effects, metrical covariates with non-linear effects, unstructured random effects, and spatial covariates, are all treated within the same general framework by assigning appropriate priors with different forms and degrees of smoothness. A main objective of this step was to see which socioeconomic factors have the most influence on the nutritional status of children and which regions are most affected by malnutrition in each country. In the second step, we apply a geoadditive latent variable model, using the three types of undernutrition as indicators of nutritional status. We have based our decision about covariates should be used in the measurement model, and which should be used in the structural equation, on the same criteria were used in section 5.3 (chapter 5).

6.3.1 Application to Childhood Malnutrition, using Separate Geoadditive Gaussian Models

In this section, we present results for the Gaussian models. The responses y_j, $j = 1,..,3$ are stunting, wasting and underweight as measurements of nutritional status. The predictor of the model considered for the analysis in this section is as follows:

$$y_{ij} = \eta_{ij} + \varepsilon_{ij} \tag{6.5}$$

$$\eta_{ij} = \beta_{0j} + f_j(Chage) + f_j(BMI) + f_j(Mageb) + f_{str}(reg) + f_{unstr}(reg) + u_i'\alpha \tag{6.6}$$

In these models, β_0 is a constant term and the covariates vector α contains all the bio-demographic and health factors which were included in the analysis of childhood disease in section 5.3 (chapter 5). The nonparametric effects are child's age, mother's BMI and mother's age at birth, which are assumed to have a nonlinear effect on the nutritional status of children in both countries, as well as the spatial effects f_{str} and f_{unstr}. The main aim in this study is to study child nutritional status by distinguishing among three response variables:

Response variables

stunting: Height-for-age, which indicates stunting.

underweight: Weight-for-age, an indication of underweight.

wasting: Weight-for-height, an indication of wasting.

The three response variables are continuous, as mentioned above. The following covariates were considered in the analysis to study child nutritional status in both countries:

Metrical covariates

Chage: Child's age in months.

BMI: Mother's body mass index.

Mageb: Mother's age at birth.

Categorical covariates

male: Child's sex : male or female (reference category).

educ: Mother's educational attainment: incomplete primary, complete primary, and incomplete secondary school; or complete secondary school and higher eduction (reference category).

trepr: Whether mother had treatment during pregnancy: yes or no (reference category).

anvis: Whether mother had antenatal care: yes or no (reference).

water: Source of drinking water: controlled water or no (reference category).

toilet: Has flush toilet at household: yes or no (reference category).

urban: Locality where respondent lives: urban or rural (reference category).

radio: Has a radio at household: yes or no (reference category).

elect: Has electricity: yes or no (reference category).

work: Mother's current working status: working or not (reference).

Spatial covariates

reg: Governorate or region where respondent resides

Results

The estimate of fixed effects of the covariates for the geoadditive Gaussian model (equation 6.2) are given in tables 6.1 through 6.3, and the nonlinear effects of child's age, mother's BMI and mother'age at birth are shown in figures 6.1 and 6.3. The regional effects are in the maps of figures 6.2 and 6.4.

Stunting

The results of the geoadditive Gaussian model show a negative relationship between male children and stunting and a positive relationship between urban area and stunting (table 6.1). These results suggest that female children in urban areas are better nourished compared to their counterparts in rural area. This finding has also been found in some previous studies in developing countries (N.B. Kandala, S. lang, and S. Klasen, 2001; Klasen, 1996; Hill and Upchurch, 1995). In addition, the educational level of the mother has a slight impact on the level of stunting and; the other categorical covariates have also either a slight or nonsignificant impact on the level of stunting in Egypt. In Nigeria, however, females whose mothers have obtained clinical care during pregnancy, have access to electricity and a flush toilet in the household are likely to be better nourished compared to their counterparts. In addition, the remaining socioeconomic factors have only a slight effect on the stunting of a child in Nigeria.

The top left panels of figures 6.1 and 6.3 display the nonparametric effect of the child's age in Egypt and Nigeria, respectively. Shown are the posterior means together with 80% and 95% pointwise credible intervals. In Egypt, we find the influence of a child's age on its nutritional status is considerably high between the age of 5 months and the age of 15 months, whilst in Nigeria this influence is increasing until the age of 20 months of age. This deterioration in nutritional status of a child begins around 5 months after birth and continuous, with an almost linear trend until the age of 15 months in Egypt and until 20 months in Nigeria. In Egypt, after 15 months of age and between the ages of 15 to 30 of months, stunting decreases, and stabilizes thereafter at a middle level. In Nigeria, on the other hand, the effect of a child's age starting from a high level in the first 20 months, declines more or less steadily until 24 to 25 months, where a bump appears in the graphics.

In looking at the mother's BMI and its impact on the level of stunting, the right panels of figures 6.1, and 6.3 show that the influence is in the form of an inverse U shape for both countries. Results for Egypt show that the mothers with BMI between 23 and 29 have a slightly higher z-score of *height-for-age* (lower stunting) measured by stunting, and the effect stabilizes at the same level thereafter. Mothers with BMI less than 20 have a lower z-score of *height-for-age*. It shows that BMI has a slight effect on the nutritional status. For Nigeria, the figure reveals that obesity of the mother is likely to pose less of a risk to the nutritional status of a child. Low BMIs of less than 18.5 suggest acute undernutrition of the mother. Furthermore, the z-score is highest at a BMI of around 30 to 35 in Nigeria and around 35 in Egypt (and thus lowest stunting).

The effect of mother's age on stunting is quite slight (third panel from top of figure 6.1 and figure 6.3. It shows that the *height-for-age* z-score is low for mother between the ages 12 to 33 years. The z-score of *height-for-age* increases (and stunting is decreasing) after age of 33 years. For Egypt, after age of 33, the effect of the mother's age stabilizes, with an almost linear trend. This effect is also witnessed in Nigerian children whose mothers are younger than 30 years of age. It shows that their children are better in their nutritional status compare to children whose mothers are in the middle age group.

Spatial effects are allocated by the model into structured and unstructured effects shown in figures 6.2 and 6.4 for Egypt and Nigeria, respectively. For Egypt, the model shows that the structured effects are significant. This indicates that the worst nutrition is implying a higher relative risk of stunting, in some cities and rural areas on the Nile Delta. Note that the unstructured effect of Egypt shows that there is no cases found in the governorates with gray color; that is because most of these areas are not populated.

For Nigeria, the data indicates that most of the regions in the southeast and some of regions in south are associated with high rates of stunted children. In other words, figure 6.4 reveals that most of the children from southeastern regions of the country (namely the states Akwa-Ibom Cross-River, Anambra, Enugu, Ebonyi, Imo and Abia) suffer from high incidences of stunting, while the northwestern region is associated with significantly low childhood undernutrition.

Wasting

Results of fixed effects parameters are shown in tables 6.3 and 6.6 for Egypt and Nigeria, respectively. In Egypt, female children whose mothers have obtained antenatal visit during their pregnancy, and have access to radio, have higher z-scores of *weight-for-height* (lower level of wasting), which implies that they are better nourished. Astonishingly, children whose mothers had clinical treatment are not associated with higher *weight-for-height*. Gender differences in childhood nutrition have been confirmed by other authors (Svedlber (1996), Klasen (1996), and Adebayo (2002)). It is found that female are better nourished than male children and the effect of gender is similar in our application. In addition, the mother's current employment status, where the mother lives (rural/urban), availability of electricity, access to controlled water, education level of mothers, availability of flush toilet at household, have either slight or statistically insignificant effects on a child's *weight-for-height* in Egypt. The results of Nigeria are unlike the case of Egypt. It is seen that only children whose mothers are currently working contribute significantly lower *weight-for-height* than their counterparts, while children whose mothers had treatment during pregnancy have no significant effect on the nutritional status. On the other side, most socioeconomic factors have a slight effect on the undernutrition status of children in Nigeria.

There is evidence that there is deterioration in a child's *weight-for-height* from the age of 5 months until the child is about 20-25 months in Egypt, where minimum z-scores of *weight-for-height* is attained and goes on to stabilize at a low level thereafter. In Nigeria, however, the deterioration increases until about 20 months (see second top left panels of figure 6.1 for Egypt and figure 6.3 for Nigeria)(see the discussion in section 4).

Figures for the mother's BMI display an almost linear trend with positive slopes. The implication is that children whose mothers have a low BMI are likely to be wasted (figure 6.1 for Egypt and figure 6.3 for Nigeria). In

Egypt, there is an almost linear fixed effect from BMI of 20 to 35.

The effect of mother's age on nutritional status of children is high for older mothers (> 30) in Egypt and it is quite similar for the underweight anthropometric indic. For Nigeria, the effect of mother's age is quite similar for the three anthropometric indices.

Figure 6.2 displays structured spatial effects (left panels) on stunting, wasting and underweight, with corresponding unstructured spatial effects (right panels), on the colored map of Egypt. The geographical panel indicates a significantly high rate of wasted children are associated with some regions in Nile Delta such as Damietta, Dakhalia and Esmaliyia.

In Nigeria, a high rates of wasted children are associated with northwestern part of the country and some districts in central Nigeria (such as Sokoto, Zemfaa, Kebbl and Kastina).

Underweight

This response variable belongs to the index of stunting and wasting. That means, a child may be underweight if s/he is either stunted, or wasted, or both. The factors associated with underweight are presented in tables 6.2 and 6.5 for Egypt and Nigeria.

The results indicate that male children were at higher risk of being underweight than female children. Children born to mothers with a secondary or higher educational level, and who obtained medical care during pregnancy, were at lower risk of malnutrition compared to the other children.

We also note that having a flush toilet in the household, whether a mother had treatment during her pregnancy are not associated with better nutritional status.

Meanwhile, the results for Nigeria seem to be more reasonable, compared to the results for Egypt. As expected and as confirmed in many previous studies, female children seem to be better nourished than male children, and the gender effect in this application is consistent with the results of these studies. Furthermore, children whose mothers are currently working, have completed at least secondary school or higher education, have made antenatal visits, have electricity, radio, or flush toilet and have access to controlled water are better nourished compared to their counterparts.

As previously mentioned, the nonlinear effects of the continuous covariates are shown in figures 6.1 and 6.3 for Egypt and Nigeria, respectively. A child may have low z-scores of *weight-of-age* if s/he is either chronically malnourished (stunted), acutely malnourished (wasted), or both. As a consequence, underweight is combination of stunting and wasting. It is obvious that the deterioration in *weight-for-age* after the first 4 or 5 months of life and low stability level was reached at age 15 months in both countries. However, and particulary in Nigeria, an improvement commenced from about 25-30 months and rise gradually until age 60 months. Hence, this seems to be an average effect of high *weight-for-age* and low *height-for-age* during this period (i.e., 25 to 60 months). The patterns of mother's BMI and mother's age at birth are quite similar to *weight-for-height*.

The bottom panels of figures 6.2 and 6.4 depict the occurrence of underweight children in both countries. They show clearly where the nutritional problems are severe. The occurrence of underweight children was particulary high in southern Nigeria, where the malnourishment problem appeared as well. In Egypt, the same regions in the Nile Delta, such as Damietta, Dakhalia and Esmaliyia which, are associated with high level of stunting and wasting, are also affected by underweight.

Variable	Mean	S.dv	10%	median	90%
const	-0.715^*	0.134	-0.894	-0.712	-0.546
male	-0.078^*	0.016	-0.098	-0.077	-0.055
urban	0.029^*	0.021	0.002	0.028	0.057
work	-0.004	0.023	-0.034	-0.005	0.025
trepr	0.001	0.027	-0.033	0.002	0.035
anvis	0.014	0.019	-0.010	0.015	0.0388
radio	-0.005	0.022	-0.033	-0.006	0.024
elect	-0.066	0.086	-0.174	-0.067	0.042
water	-0.004	0.026	-0.037	-0.004	0.029
educ	0.017	0.022	-0.010	0.018	0.0455
toilet	-0.028	0.040	-0.077	-0.027	0.0208

Table 6.1: Fixed effects on stunting in Egypt.

Variable	Mean	S.dv	10%	median	90%
const	-0.451^*	0.123	-0.610	-0.453	-0.290
male	-0.0757^*	0.013	-0.093	-0.075	-0.058
urban	0.005	0.015	-0.014	0.005	0.025
work	0.005	0.0173	-0.016	0.005	0.028
trepr	-0.025	0.0215	-0.053	-0.026	0.0045
anvis	0.034^*	0.014	0.017	0.034	0.051
radio	0.017	0.017	-0.005	0.017	0.038
elect	-0.071	0.066	-0.158	-0.074	0.015
water	-0.006	0.020	-0.030	-0.006	0.017
educ	0.029^*	0.015	0.010	0.029	0.050
toilet	0.0183	0.032	-0.024	0.017	0.059

Table 6.2: Fixed effects on underweight in Egypt.

Variable	Mean	S.dv	10%	median	90%
const	0.038*	0.132	-0.132	0.040	0.203
male	−0.055*	0.014	-0.074	-0.056	-0.037
urban	-0.014	0.018	-0.038	-0.013	0.007
work	0.01	0.019	-0.015	0.010	0.035
trepr	−0.030*	0.023	-0.061	-0.030	-0.0005
anvis	0.036*	0.0163	0.0146	0.036	0.057
radio	0.028*	0.019	0.004	0.028	0.053
elect	-0.043	0.072	-0.142	-0.041	0.043
water	-0.007	0.021	-0.036	-0.007	0.019
educ	0.020	0.017	-0.001	0.020	0.045
toilet	-0.051	0.035	-0.123	-0.050	0.017

Table 6.3: Fixed effects on wasting in Egypt.

6.3.2 Analyses using Latent Variable Models for Continuous Responses

In this section, our interest is in analyzing the three types of undernutrition status of children in both countries using latent variable models, and in investigating how they can be established as indicators of the latent variable "undernutrition status". Based on the previous separate analyses, we are able to determine which factors can have direct effects and which can have indirect effects on the indicators. The analysis begins with two major parts, corresponding to continuous outcomes versus mixed outcomes later on. Within each part, conventional analysis using continuous latent variables will be described first, followed by recent extensions that add binary indicators of childhood disease to the analysis with latent variables (see chapter 7).

We start using the easiest model possible, a classic factor analysis for continuous indicators. The predictor of the structural equation of the model yields LMV0:

$$\eta = 0 \tag{6.7}$$

Estimates of factor loadings are depicted in table 6.7. The estimated mean

Variable	Mean	S.dv	10%	median	90%
const	-1.133^{*}	0.154	-1.33	-1.133	-0.94
male	-0.117^{*}	0.030	-0.156	-0.117	-0.077
urban	0.032	0.039	-0.020	0.032	0.083
work	0.027	0.033	-0.016	0.025	0.070
trepr	0.075^{*}	0.039	0.026	0.074	0.128
anvis	0.147^{*}	0.039	0.095	0.147	0.199
radio	0.017	0.037	-0.030	0.017	0.063
elect	0.131^{*}	0.039	0.077	0.129	0.180
water	0.044	0.044	-0.008	0.043	0.106
educ	-0.543	0.943	-1.766	-0.509	0.606
toilet	0.078^{*}	0.048	0.013	0.078	0.140

Table 6.4: Fixed effects on stunting in Nigeria.

Variable	Mean	S.dv	10%	median	90%
const	-0.710^{*}	0.121	-0.863	-0.718	-0.551
male	-0.032^{*}	0.022	-0.061	-0.033	-0.004
urban	-0.022	0.030	-0.059	-0.023	0.016
work	0.044^{*}	0.026	0.009	0.043	0.077
trepr	0.014	0.031	-0.027	0.014	0.053
anvis	0.079^{*}	0.030	0.040	0.080	0.116
radio	0.035^{*}	0.028	0.0007	0.034	0.072
elect	0.065^{*}	0.029	0.024	0.067	0.101
water	0.046^{*}	0.033	0.001	0.047	0.089
educ	0.063^{*}	0.038	0.013	0.064	0.111
toilet	0.105^{*}	0.044	0.051	0.106	0.159

Table 6.5: Fixed effects on underweight in Nigeria.

Variable	Mean	S.dv	10%	median	90%
const	-0.041^*	0.127	-0.214	-0.032	0.116
male	0.026	0.024	-0.005	0.025	0.058
urban	-0.051	0.030	-0.111	-0.050	0.006
work	0.049^*	0.027	0.011	0.050	0.083
trepr	-0.046	0.034	-0.116	-0.045	0.022
anvis	-0.038	0.030	-0.076	-0.039	0.0004
radio	0.018	0.032	-0.023	0.020	0.060
elect	-0.019	0.031	-0.060	-0.019	0.020
water	0.028	0.036	-0.016	0.026	0.075
educ	0.030	0.041	-0.022	0.030	0.0847
toilet	0.0037	0.048	-0.055	0.0006	0.068

Table 6.6: Fixed effects on wasting in Nigeria.

Parameter	Mean	Std	2.5%	97.5%
	Factor Loadings			
1.stuntingλ_{11}	0.72	0.0149	0.69	0.74
2.underweightλ_{21}	1.053	0.0079	1.04	1.06
3.wastingλ_{31}	0.757	0.0123	0.734	0.781

Table 6.7: Results of Model LVM0 of Z-scores indicators for Egypt with $\eta = 0$.

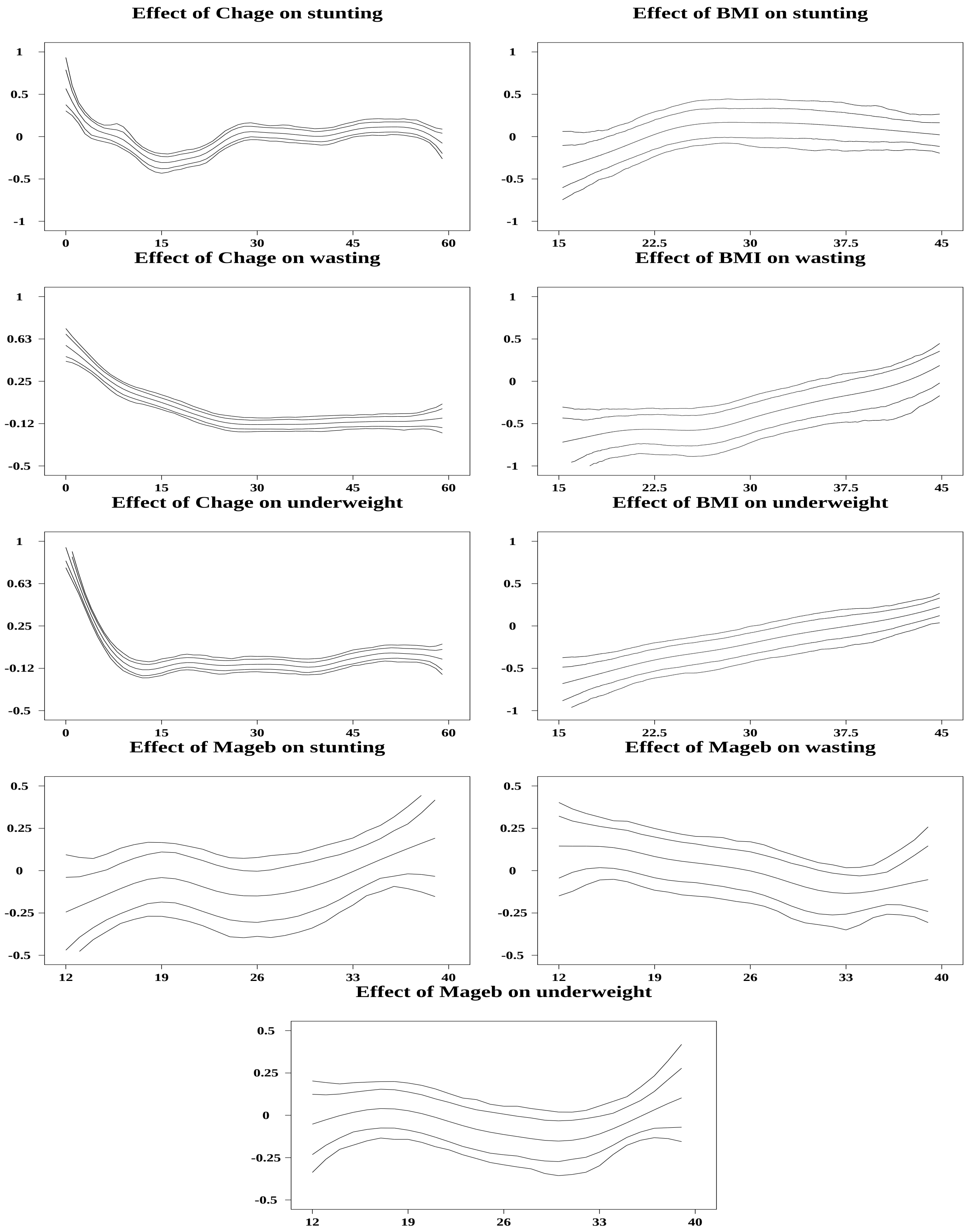


Figure 6.1: Posterior means of nonparametric effects in stunting (top), wasting (second from top) and underweight (bottom) for child's age (left top to third from top), mother's BMI (right top to third from top) and mother's age (last three panel from bottom) for Gaussian semiparametric model in Egypt.

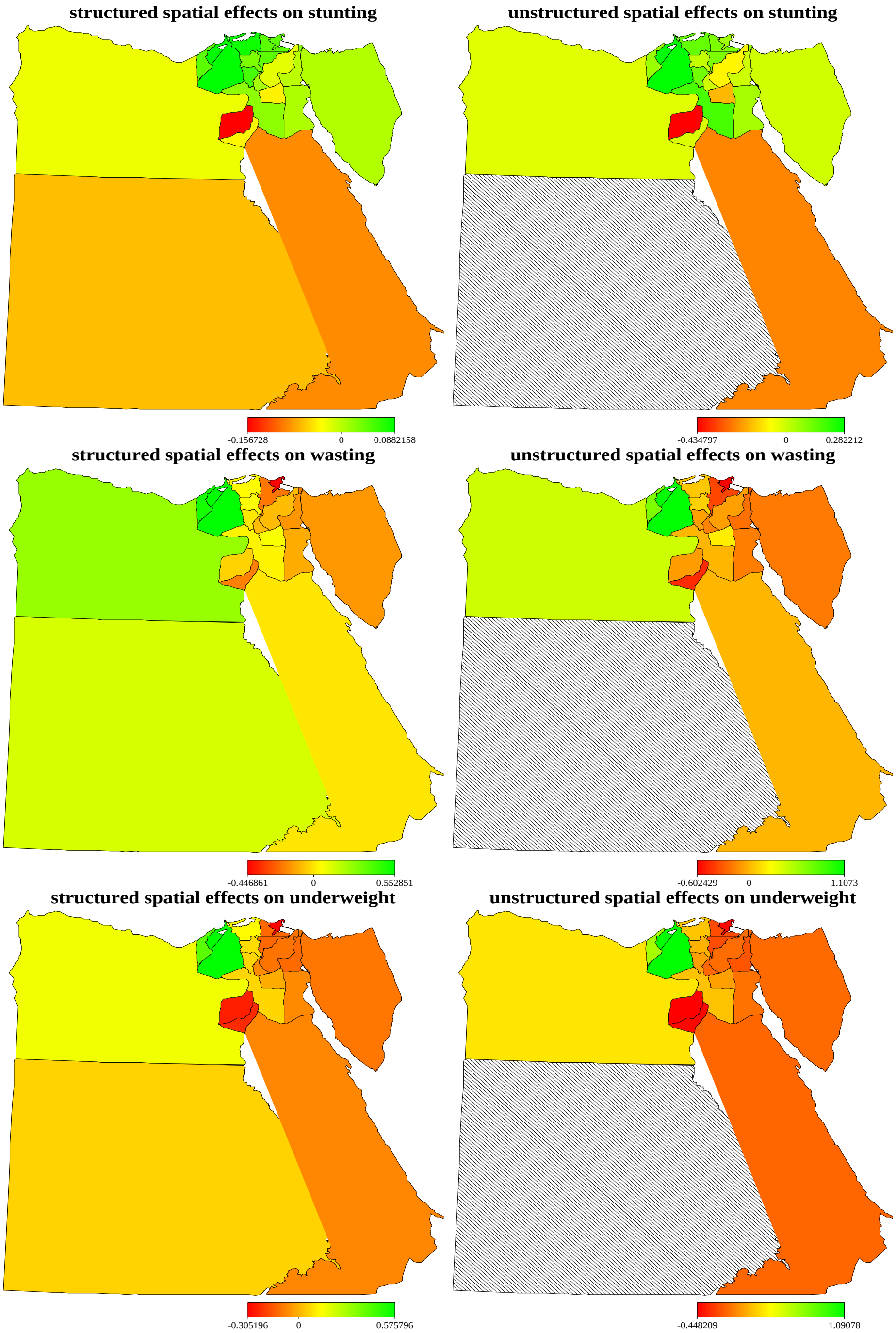


Figure 6.2: Colored maps of Egypt, showing posterior means of structured (left) and unstructured (right) spatial effects in stunting (top), wasting (middle) and underweight (bottom) for Gaussian semiparametric.

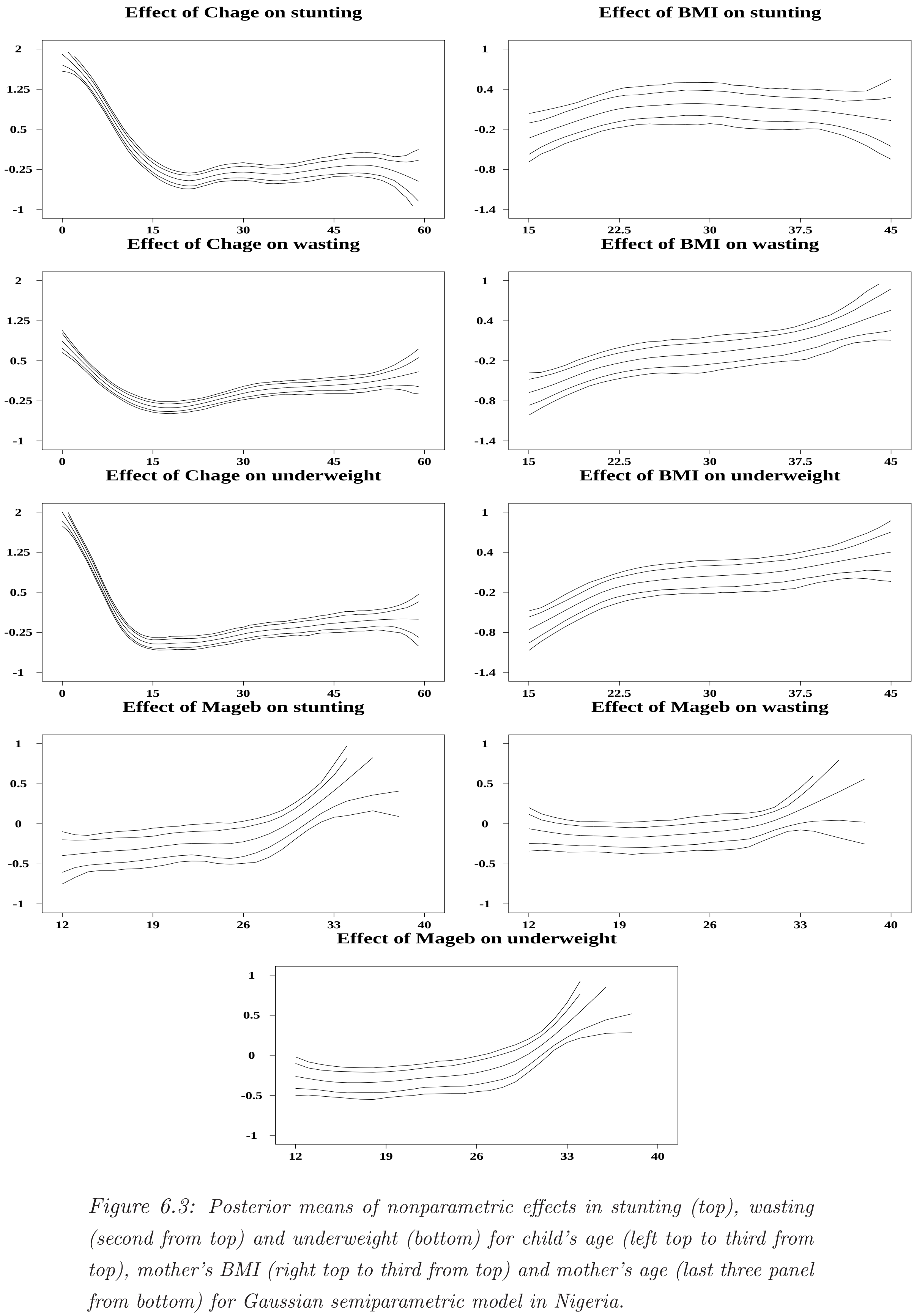


Figure 6.3: Posterior means of nonparametric effects in stunting (top), wasting (second from top) and underweight (bottom) for child's age (left top to third from top), mother's BMI (right top to third from top) and mother's age (last three panel from bottom) for Gaussian semiparametric model in Nigeria.

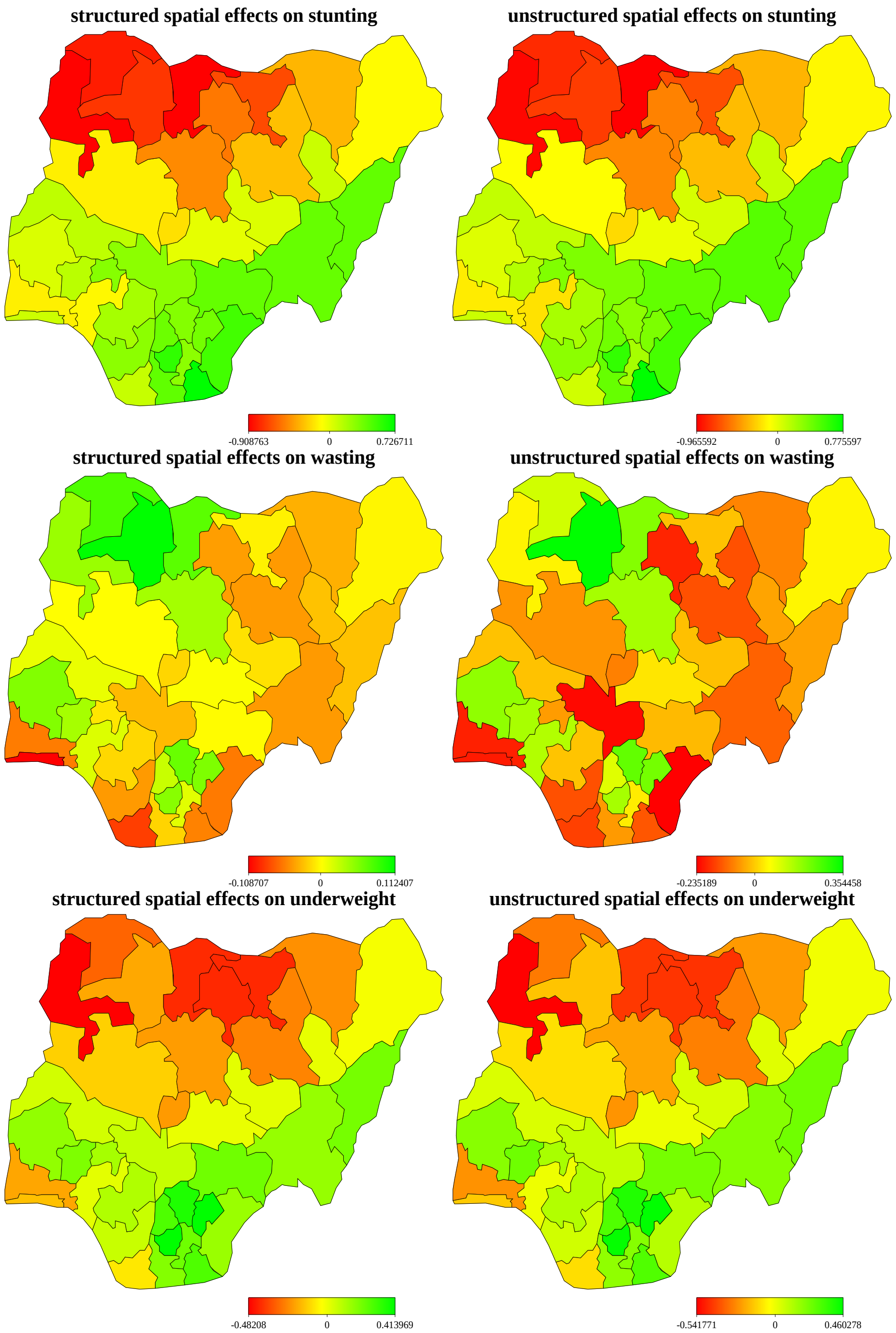


Figure 6.4: Colored maps of Nigeria, showing posterior means of structured (left) and unstructured (right) spatial effects in stunting (top), wasting (middle) and underweight (bottom) Gaussian semiparametric model.

factor loadings show that indicator 2 (*weight-for-age*) has the highest factor loading.

The classic factor analysis model has been extended by introducing direct and indirect parametric covariates, which modified the latent construct.

The next model was selected based on the previous separate analyses. This leads to the latent variable model (equation 6.6).

In the fundamental analysis of Egypt (LVM1), the vector a_j (equation 6.3) comprises the covariates urban, mother working, treatment during pregnancy, educational level of mothers, access to flush toilet, and availability of electricity, with direct effects on y_j; and u_i' comprises the remaining categorical covariates sex, antenatal visits and access to controlled water, having common effects on the latent variable υ. In Nigeria, the covariates which are included in the indirect effects at the beginning were child's sex, employment status of mother, treatment during pregnancy, availability of radio and access to controlled water. Nonparametric functions and the spatial effects were included in the geoadditive predictor for υ.

The results of geoadditive latent variable models for Egypt are shown in table 6.8 for model LVM1. They indicate that female children whose mothers obtained antenatal visits during pregnancy are more likely to be better nourished, and these factors have significant effect on the latent variable (nutritional status). But on the other hand, the access to controlled water has a slight effect on the nutritional status of children in Egypt. Regarding parametric direct effects, the results indicate a significant effect of urban location, education level of mothers, treatment during her pregnancy, availability of electricity, radio and type of toilet on the factor loading of *weight-for-age*(λ_{21}). While the covariates of radio, treatment during pregnancy and type of toilet, urban (at 90%) and electricity (at 90% confidence intervals) have significant effect on *weight-for-height*. Effects of toilet, electricity and place of residence are, however, negative on the indicators y_2 (underweight) and y_3 (wasting). The cause of these negative signs could be due to the following reasons: First, in the analysis of latent models, we used three indicators (which were assumed to have high level of correlations among each other) instead of one indicator, which was used by the separate analysis.

Because of that, it is difficult to compare the results of the LVM with the previous separate analysis.

Second, we have found that 66.4% of the households that have access to electricity and flush toilet are located in the rural areas and 33.7% of the households are located in the urban areas, therefore the corresponding effects of the rural areas (which were assumed to be reference category) are higher than their counterparts in the urban areas.

Third, it is observed that the indicators have a higher correlation which can affect the results, so we have made a further analysis excluding the indicator of wasting (*weight-for-hight*) to examine the effects of various factors on the other indicators (underweight and stunting), and results are compared with analysis when all three indicators are present (see table 6.14). Moreover, the level of the education of the mother is the only covariate which has a significant effect on y_1. Note that some variables, such as radio and electricity, were associated with nonsignificant effect on the *weight-for-age* in the separate analysis, and they become significant on the second indicator y_2 as shown by model LVM1. Still, the difference is not large between the results of model LVM1 and the previous results that were obtained using Gaussian geoadditive models. For further analysis, we included the parametric direct covariates which were insignificant in LVM1 in the indirect parametric effects of the model LVM2, and they still seem to be insignificant or have slight effects on the nutritional status of children in Egypt (see table 6.9).

The results of the further analysis using only two indicators (stunting and underweight) show that the child's sex and antenatal visits have significant indirect effects on the nutritional status by LVM3 in Egypt (table 6.13). On the other hand, the results for the covariates of direct effects are yet more reasonable compared to the results by LVM1 and LVM2 (tables 6.8 and 6.9). It shows that education level of mothers are associated with higher *height-for-age* (thus lowest level of stunting) and with higher *weight-for-age* (thus lowest level of underweight). Furthermore, the results indicate that the variable flush toilet has a negative effect on the indicator of underweight.

The factor loadings estimates show that *weight-for-age* is more serious (due to its high factor loading of 0.968) in Egypt compared to its reference population. With regards to nonlinear effects on the latent variable nutritional status, their patterns are similar to the patterns of the separate analysis, as can be seen in figure 6.5, and in addition, the patterns of LVM1 and LVM2 are quite similar. The nonlinear effect of child's age (top panels of figure 6.9) shows that the deterioration sets in right after 5 months of birth and continues, until 20-25 months; then the age impact declines. The second panel for figure 6.9) shows that the children of mothers who have BMI less than 20 are more likely to be undernourished. In addition, the effect of mother's age is quite similar to the previous analysis.

The spatial effect is displayed in figure 6.6 and it shows that some rural areas located in the Nile Delta are associated with a high rate of undernutrition; as mentioned in some previous studies of undernutrition, this part of the country is associated with a high rate of illness among children under 5 (see results of childhood disease, chapters 3 and 5) which lead to problems in the nutritional status of children in these areas. Furthermore, figure 6.11 indicates that significant high undernutrition rates associated with the northeast areas. These results indicate that children living in the northeast areas, are more affected by stunting and wasting. This finding is not surprising and it is consistent with the results of the analysis of childhood disease in chapter 3 and chapter 5.

Parametrer	Mean	Std	2.5%	10%	90%	97.5%
		Factor Loadings				
stunting λ_{11}	0.66**	0.014	0.636	0.673	0.681	0.692
underweight λ_{21}	0.982**	0.005	0.974	0.976	0.989	0.996
wasting λ_{31}	0.712**	0.011	0.691	0.698	0.727	0.736
		Parametric Indirect Effects				
male	−0.116**	0.039	-0.192	-016	-0.066	-0.038
anvis	0.117**	0.040	0.039	0.066	0.169	0.197
water	0.123	0.101	-0.074	-0.006	0.253	0.321
		Parametric Direct Effects				
urban(a_{11})	0.031	0.032	-0.03	-0.009	0.072	0.094
work(a_{12})	-0.004	0.038	-0.080	-0.052	0.045	0.073
trepr(a_{13})	-0.042	0.049	-0.140	-0.106	0.019	0.050
elect (a_{14})	-0.165	0.151	-0.469	-0.360	0.029	0.126
radio(a_{15})	0.027	0.038	-0.048	-0.022	0.077	0.103
educ(a_{16})	0.063**	0.023	0.0184	0.033	0.091	0.108
toilet(a_{17})	-0.061	0.072	-0.200	-0.154	0.029	0.076
urban(a_{21})	−0.006*	0.005	-0.019	-0.014	-0.0009	0.004
work(a_{22})	-0.003	0.012	-0.031	-0.024	0.010	0.016
trepr(a_{23})	−0.076**	0.0175	-0.109	-0.103	-0.055	-0.048
elect(a_{24})	−0.192**	0.075	-0.368	-0.337	-0.117	-0.105
radio(a_{25})	0.075**	0.012	0.0560	0.0608	0.090	0.109
educ(a_{26})	0.047**	0.0068	0.027	0.0369	0.053	0.056
toilet(a_{27})	−0.1740**	0.018	-0.204	-0.197	-0.146	-0.136
urban(a_{31})	−0.035*	0.025	-0.084	-0.067	-0.001	0.013
work(a_{32})	-0.01	0.031	-0.071	-0.050	0.030	0.050
trepr(a_{33})	−0.056*	0.039	-0.132	-0.107	-0.005	0.022
elect(a_{34})	-0.120	0.127	-0.376	-0.290	0.043	0.121
radio(a_{35})	0.077**	0.032	0.016	0.036	0.118	0.140
educ(a_{36})	0.011	0.018	-0.026	-0.012	0.034	0.047
toilet(a_{37})	−0.154**	0.058	-0.268	-0.228	-0.079	-0.037
		Smoothing Parameters				
Chage	0.014**	0.011	0.003	0.005	0.026	0.042
BMI	0.002**	0.004	0.0003	0.0005	0.005	0.013
Mageb	0.003**	0.004	0.0004	0.0006	0.006	0.012
reg	0.570**	0.242	0.026	0.33	0.867	1.189

Table 6.8: Results of LVM1, including direct and indirect effects in Egypt.
(**: Statistically significant at 2.5% and 10%)

Parametrer	Mean	Std	2.5%	10%	90%	97.5%
		Factor Loadings				
stunting λ_{11}	0.648**	0.013	0.621	0.631	0.665	0.675
underweight λ_{21}	0.959**	0.003	0.955	0.956	0.965	0.967
wasting λ_{31}	0.696**	0.010	0.675	0.682	0.710	0.717
		Parametric Indirect Effects				
male	-0.162**	0.027	-0.217	-0.196	-0.127	-0.108
anvis	0.085**	0.030	0.026	0.046	0.125	0.144
work	0.020	0.035	-0.048	-0.024	0.067	0.090
trepr	-0.055	0.044	-0.140	-0.114	0.001	0.031
elect	-0.252*	0.136	-0.521	-0.427	-0.075	0.012
radio	0.034	0.035	-0.035	-0.012	0.079	0.102
		Parametric Direct Effects				
water(a_{11})	0.032	0.044	-0.054	-0.023	0.088	0.116
educ(a_{12})	0.043*	0.022	-0.00059	0.014	0.071	0.087
toilet(a_{13})	-0.009	0.072	-0.150	-0.105	0.084	0.131
urban(a_{14})	-0.010	0.036	-0.081	-0.0579	0.037	0.061
water(a21)	0.054**	0.013	0.033	0.0408	0.069	0.093
educ (a_{22})	0.0214**	0.006	0.0087	0.001	0.027	0.033
toilet(a_{23})	-0.085**	0.030	-0.127	-0.120	-0.038	-0.028
urban(a_{24})	-0.064**	0.022	-0.092	-0.088	-0.030	-0.024
water(a_{31})	0.040	0.035	-0.029	-0.005	0.085	0.108
educ (a_{32})	-0.006	0.018	-0.043	-0.0307	0.016	0.029
toilet(a_{33})	-0.082*	0.061	-0.199	-0.160	-0.003	0.039
urban (a_{34})	-0.075**	0.030	-0.135	-0.115	-0.035	-0.015
		Smoothing Parameters				
Chage	0.015**	0.012	0.004	0.0057	0.027	0.045
BMI	0.002**	0.002	0.0003	0.0004	0.005	0.009
Mageb	0.002**	0.0035	0.0003	0.0006	0.006	0.012
reg	0.596**	0.249	0.276	0.341	0.918	1.238

Table 6.9: Results of LVM2, including direct and indirect effects in Egypt.

Table 6.10 displays the estimation of the factor loadings in the case of Nigeria. It shows that the indicator *weight-for-age* has the highest factor loadings. That means the most effect on the z-scores is on underweight for age and is followed by the indicator of stunting.

Table 6.11 provides various covariates that are used in model LM1 based on the previous (separate analyses) results. Gender differences in malnutrition are most pronounced at young ages. Girls are significantly more likely than similarly aged boys to be better nourished. The effect of having treatment during pregnancy seems to be nonsignificant on Z-scores "malnutrition status". Availability of radio and access to controlled water have

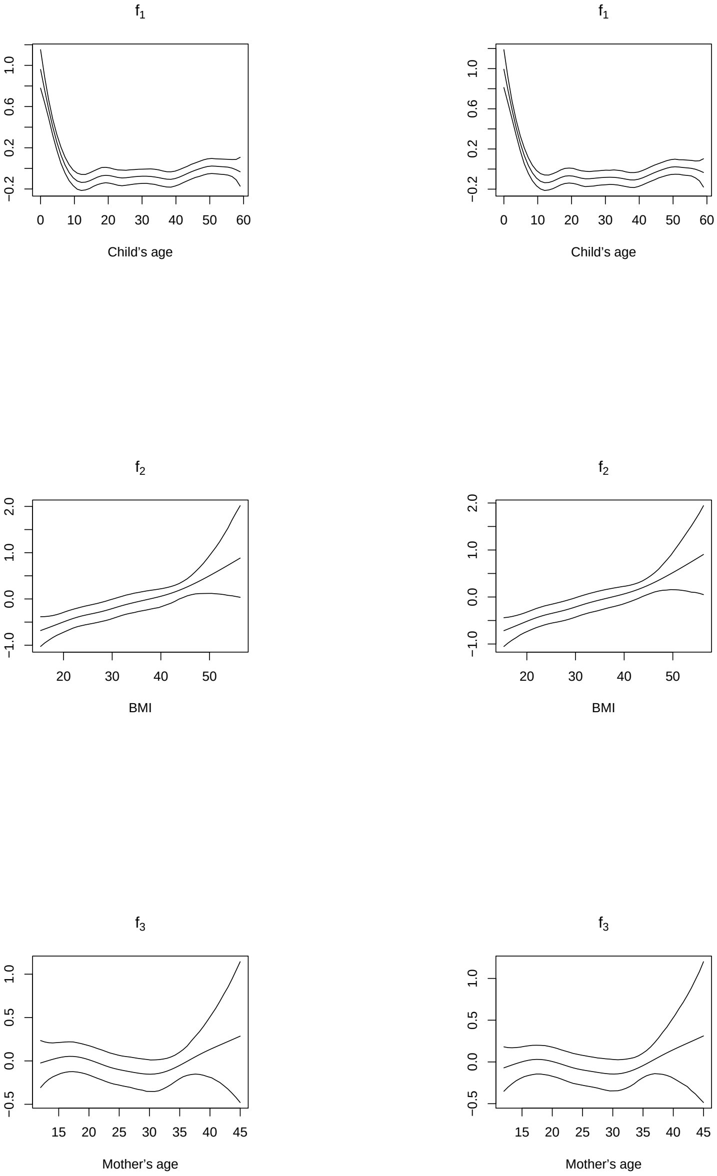


Figure 6.5: Non-linear effects from top to bottom: child's age, mother's BMI and mother's age at birth (for model LVM1-left panels), child's age, mother's BMI and mother's age at birth (for model LVM2-right panels) on the indicators of a latent variable "malnutrition status" of children for Egypt, using Bayesian latent variable models for continuous responses.

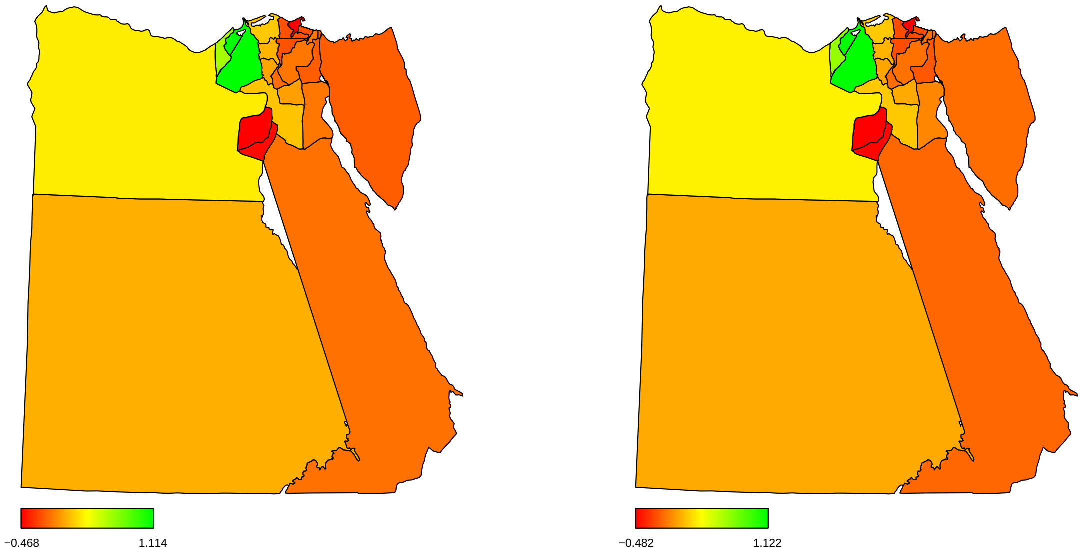


Figure 6.6: Posterior mean for latent variable model for LVM1 (left panel) and LVM2 (right panel) on malnutrition status for Egypt.

Parametrer	Mean	Std	2.5%	97.5%
	Factor Loadings			
1.stunting λ_{11}	1.244	0.020	1.206	1.286
2.underweight λ_{21}	1.3651	0.008	1.353	1.383
3.wasting λ_{31}	0.770	0.015	0.739	0.801

Table 6.10: Results of Model LVM0 of Z-scores indicators in Nigeria with $\eta = 0$.

Parametrer	Mean	Std	2.5%	10%	90%	97.5%
		Factor Loadings				
stunting λ_{11}	1.037**	0.021	0.994	1.009	1.064	1.077
underweight λ_{21}	1.217**	0.0038	1.207	1.219	1.220	1.221
wasting λ_{31}	0.7102**	0.017	0.675	0.687	0.732	0.744
		Parametric Indirect Effects				
male	-0.0074*	0.052	-0.169	-0.137	-0.003	0.032
work	0.046	0.063	-0.079	-0.034	0.128	0.172
trepr	0.015	0.053	-0.089	-0.0517	0.083	0.121
radio	0.021*	0.027	0.0320	-0.014	0.056	0.075
water	0.060*	0.032	-0.003	0.016	0.102	0.123
		Parametric Direct Effects				
urban(a_{11})	0.0016	0.059	-0.113	-0.075	0.079	0.118
anvis(a_{12})	0.515**	0.059	0.039	0.044	0.592	0.629
toilet(a_{13})	0.155**	0.042	0.072	0.100	0.209	0.238
elect(a_{14})	0.266**	0.059	0.151	0.192	0.341	0.383
educ(a_{15})	0.041	0.040	-0.036	-0.009	0.095	0.121
urban(a_{21})	-0.128**	0.010	-0.155	-0.141	-0.117	-0.110
anvis(a_{22})	0.184**	0.016	0.164	0.165	0.207	0.212
toilet(a_{23})	0.065**	0.009	0.047	0.055	0.078	0.089
elect(a_{24})	0.162**	0.007	0.146	0.152	0.171	0.176
educ(a_{25})	-0.009	0.012	-0.0395	-0.0324	0.003	0.043
urban(a_{31})	-0.176**	0.048	-0.273	-0.239	-0.114	-0.082
anvis(a_{32})	-0.177**	0.048	-0.272	-0.239	-0.115	-0.084
toilet(a_{33})	-0.033	0.035	-0.103	-0.079	0.012	0.036
elect (a_{34})	0.003	0.048	-0.094	-0.059	0.066	0.098
educ(a_{35})	-0.032	0.033	-0.097	-0.075	0.011	0.033
		Smoothing Parameters				
Chage	0.0343**	0.027	0.008	0.011	0.063	0.104
BMI	0.003**	0.004	0.0006	0.001	0.008	0.014
Mageb	0.003**	0.004	0.0004	0.0006	0.006	0.014
reg	0.108**	0.041	0.048	0.063	0.162	0.207

Table 6.11: Results of LVM1, including direct and indirect effects in Nigeria.

a significant effect on child malnutrition, within the confidence interval of 97%. Although there is some evidence of the relationship between child nutrition and mother's working status, it is seen to be nonsignificant as to the indirect effect of childhood malnutrition in LVM1, but it has a significant effect on the indicator of *age-for-weight* in LVM2 (table 6.12). Using well water was strongly associated with higher Z-scores (lowest undernutrition), particularly in urban areas. The results of parametric direct effect are quite consistent with the separate analysis of Nigeria, though some covariates seem to have a negative coefficient such as urban (a_{21}), (a_{31}) and antenatal visit (a_{32}), which could be resulted of the inclusion of three indicators instead of one indicator in the separate analysis.

Model LVM1 has been extended or changed to model LVM2 by including some covariates that have direct effects to the parametric direct covariates in LVM2. The results of model LVM2 (table 6.12) show that most of the parametric direct covariates are significant and remained quite stable when including these covariates in the direct parametric effects. It demonstrates that the female children whose mothers are educated, had treatment during their pregnancy, have an access to controlled water, have an access to radio and working currently have higher Z-score of *(weight-for-age)* and are better nourished. However, males whose mothers currently working are associated with a higher level of *(weight-for-height)*(at 97%). Although working status has a slight effect on the indicator of stunting, it is associated with other indicators. According to the covariate of radio, it has mostly nonsignificant effect. Moreover, the results of LVM2 indicate a negative effect of the education on the indicator 2.

The results of LVM3 (table 6.14) indicate that the antenatal visits and the availability of electricity are associate positively with nutritional status. Regards the direct covaraites, the females and the education level of mothers have a positive significant effect on the indicator of stunting. While, only the work status is associated positively with the indicator of the underweight. The factor loadings estimates show that the *weight-for-height* is not serious compared to the case of Egypt, however, the stunting seem to be more serious in Nigeria (its higher factor loading of 1.14).

Figure 6.5 (top panels) show the non-linear effect of the child'age to be associated with a malnutrition status in Nigeria for LVM1 and LVM2, respectively. It shows that the rates of malnutrition of children increases sharply from about 5 to around 20 months of age. The rates of malnutrition are at low level between 20 and 30 months of age, then rise again through the remainder of the third year. This pattern highlights the first two years of life as the most nutritionally vulnerable for children in Nigeria.

The second panels of figure 6.5 show the nonlinear effect of the BMI of the mother. It seems quite reasonable, because the obesity of the mother probably poses the less of a risk for the child's nutritional status, due to the fact that a very low BMI suggested acute undernutrition of the mother. The Z-score is highest (and thus stunting lowest) at a BMI of around 30-40 months.

The effect of mother's age seems to be slight on the Z-scores of children up till about the age of 25 months; thereafter, there is a strong effect shown. In addition, the patterns of the nonlinear effect in LVM3 (figure 6.10) are similar to the patterns of LVM1 and LVM2.

Figures 6.8 and 6.12 show that the districts in the southeastern through the southern part of the country are associated with better nutrition of children in Nigeria.

6.4 Discussion

The results of estimating the separate Gaussian models (set out in equation 6.6) and from estimating the geoadditive latent variable models with continuous response variables (set out in equation 5.17) are indicated and suggest the following:

Child's sex

The likelihood of being stunted and underweight was lower for girls than for boys; a finding consistent with Sevelberg (1996), Klasen (1996), Lavy et. al (1996), Kandala et al. (2001a) and Kandala et. al (2001b), Adebayo

Parametrer	Mean	Std	2.5%	10%	90%	97.5%
Factor Loadings						
stunting λ_{11}	1.041**	0.021	1.00	1.02	1.079	1.095
underweight λ_{21}	1.191**	0.007	1.178	1.187	1.208	1.210
wasting λ_{31}	0.673**	0.017	0.644	0.656	0.703	0.714
Parametric Indirect Effects						
urban	-0.057	0.049	-0.153	-0.119	0.011	0.044
anvis	0.054	0.065	-0.058	-0.013	0.153	0.198
toilet	0.142**	0.059	0.017	0.060	0.212	0.250
elect	0.0683*	0.056	-0.026	0.010	0.151	0.186
Parametric Direct Effects						
male(a_{11})	-0.238**	0.0518	-0.321	-0.285	-0.153	-0.119
work(a_{12})	0.09	0.055	-0.042	-0.007	0.134	0.168
trepr(a_{13})	0.155*	0.069	-0.004	0.041	0.226	0.274
water(a_{14})	0.083**	0.035	0.0148	0.0384	0.127	0.153
educ(a_{15})	0.216**	0.039	0.143	0.167	0.265	0.291
radio(a_{16})	0.062	0.0300	-0.029	-0.0095	0.0711	0.093
male(a_{21})	-0.064**	0.0138	-0.082	-0.067	-0.032	-0.030
work(a_{22})	0.109**	0.0176	0.051	0.056	0.085	0.107
trepr(a_{23})	0.072**	0.023	0.024	0.026	0.085	0.117
water(a_{24})	0.048**	0.007	0.039	0.043	0.057	0.065
educ (a_{25})	0.067**	0.013	0.0507	0.058	0.074	0.076
radio(a_{26})	0.047**	0.0056	0.004	0.005	0.020	0.039
male(a_{31})	0.051*	0.042	-0.015	0.010	0.119	0.148
work(a_{32})	0.096*	0.0453	-0.006	0.021	0.135	0.163
trepr(a_{33})	-0.056	0.056	-0.182	-0.141	0.005	0.045
water (a_{34})	0.001	0.028	-0.054	-0.036	0.037	0.056
educ (a_{35})	-0.076**	0.032	-0.135	-0.115	-0.035	-0.015
radio(a_{36})	0.0018	0.0248	-0.068	-0.050	0.013	0.032
Smoothing Parameters						
Chage	0.035**	0.028	0.008	0.01	0.065	0.107
BMI	0.004**	0.0056	0.0006	0.001	0.010	0.018
Mageb	0.003**	0.0045	0.0004	0.0006	0.007	0.015
reg	0.121**	0.045	0.055	0.071	0.175	0.227

Table 6.12: Results of LVM2, including direct and indirect effects in Nigeria.

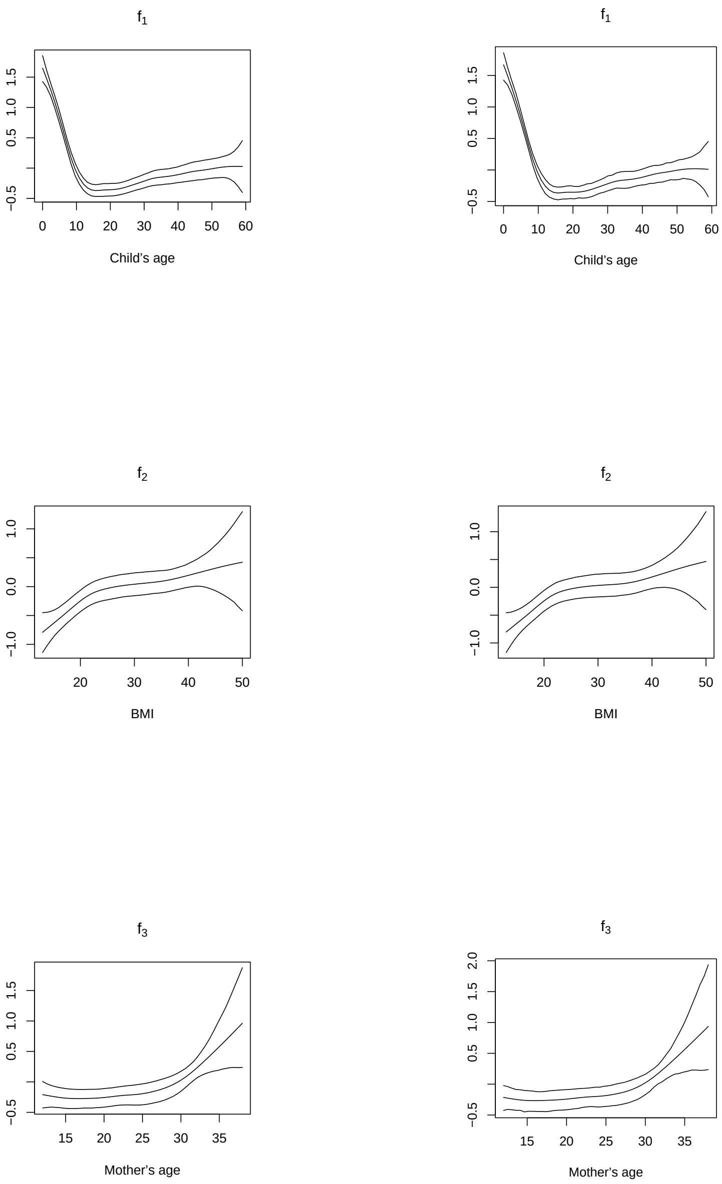


Figure 6.7: Non-linear effects from top to bottom: child's age, mother's BMI and mother's age at birth (for model LVM1-left panels), child's age , mother's BMI and mother's age at birth (for model LVM2-right panels) on the indicators of a latent variable "malnutrition status" of children for Nigeria using Bayesian latent variable model for continuous responses.

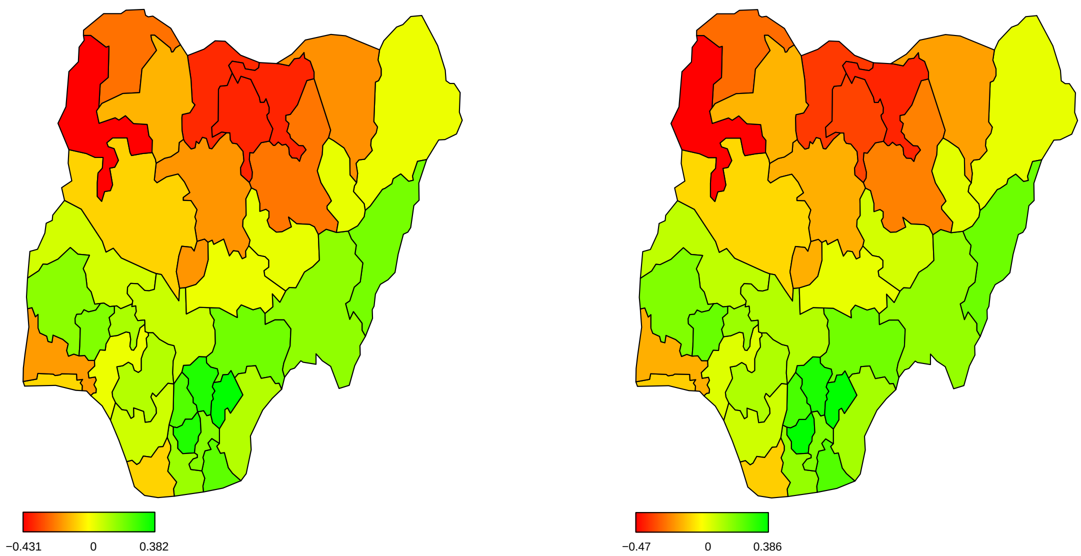


Figure 6.8: Posterior mean for latent variable model for LVM1 (left panel) and LVM2 (right panel) on malnutrition status for Nigeria.

(2002), Borohooah (2002); on the other hand, Gibson (2001) did not find any significant gender difference between the *height-for-age* and the *weight-for-age* in Papua, New Guinea.

Stunting, underweight and wasting among children by residence

Only in Egypt are urban children less likely than their rural counterparts to be stunted, as shown in the separate analysis, where the quality of health environments and sanitation are found in urban areas and these results are reasonable. On the other hand, although, rural living was expected to have many problems, such as poor health, use of unprotected water supplies, lack of charcoal as fuel, lack of milk consumption, and lack of personal hygiene, which were the risk factors for stunting, wasting and underweight, the results for both countries indicate that the place of residence is not associated with significant effects on wasting and underweight, and for stunting in the case of Nigeria (see subsection 6.3.1). This is consistent with some studies, but not with others: Adebayo (2002) found that where the mother lives (rural/urban) has no statistical significance for child's *weight-for-height*, and a similar impact of where the mother lives, as in *height-for-age*, is observed in *weight-for-age*, though Kandala found that urban ares have a statistical significance for a child's *height-for-age* in Tanzania and Malawi (see also Lavy et. al. 1996, Gibson 2001, and Borooah, 2002).

Mother's education

Maternal education, which is related to household wealth, is a determinant of good child-care knowledge and practices. In Nigeria, the education attainment of mothers is mostly significant in the analysis of LVM, (as well as in Egypt) and it has a significant effect on the underweight of a child in both countries in separate analyses, and it reduced the likelihood of children being malnourished. The results with two indicators are quite similar to the results with three indicators with regard to this variable. This result supports the suggestion that an educated mother assumes the responsibility of taking a sick child to receive care health. Further, the time that mothers spend discussing their child's illness with a doctor is almost directly proportional to their level of education: in consequence, illiterate women (and their sick children) get much less out of visiting a doctor than do literate women. These findings are consistent with many studies in the context of developing countries (Africa Nutrition chartbooks 1996, Borooah 2002, Larra et. al., 2004), which reported that maternal education has a strong and significant effect on stunting. They found that, at low levels of education, effects on stunting are small or negligible, and they increase only at secondary or higher levels. On the other hand, the result is the opposite of the cases of some developing countries which are in Latin America and Andean countries.

Mother who is currently working

Work has a positive and significant effect on the malnutrition status (wasting and underweight) of children in Nigeria; however it turned into a nonsignificant effect on the children's undernutrition status in the analysis of LVM for Nigeria as well as in the results for Egypt in both analyses. This suggests that wasting and underweight are low among children of mothers who are currently working in Nigeria. On the other hand, in Egypt, the fact that nutritional status of children of all socioeconomic levels, as represented by current employment of mothers, suggests that insufficient food intake may be not affected by the current working status of mothers. The results are consistent with some previous studies and not consistent with others. Some studies reported that when mothers are working, the household income is increased and the access to better food will be increased, as well as the access to a quality level of medical care. On the other hand, and as mentioned previously in the analysis of childhood disease, when mothers are employed outside the home, curtails the duration of full breastfeeding and necessitates supplementary feeding, usually by illiterate care-takers, which might affect the health of children negatively.

Source of drinking water

In Egypt, the results indicate that the source of water has no statistical significant effect on child nutritional status. This suggests that socioeconomic differences, represented by source of water, can not fully explain the level of stunting, underweight and wasting in Egypt. However, a household's source of drinking water is associated with the nutritional status of a child in Nigeria (*weight-for-age*) in separate analysis, and it seems to be mostly significant in the results of LVM. In other words, the source of water is associated with the nutritional status of a child through its impact on the risk of childhood diseases such as diarrhea, and is affected indirectly as a measure of wealth and availability of water.

Type of toilet

The type of toilet used by a household is an indicator of household wealth and a determinant of environmental sanitation. This means that, poor households, which are mostly located in rural areas for both countries, are less likely to have sanitary toilet facilities. In consequence, these results in increased risk of childhood diseases, which contribute to malnutrition. Regarding Egypt, it seems to have a nonsignificant statistical effect on the nutrition status of children in separate analysis, and have negative significant effects in LVM analysis even when we used two indicators instead of three. For Nigeria, the results indicate that in households where a flush toilet exits, stunting, underweight (separate analysis) are significantly lower and the nutritional status on children (analysis with LVM) is better.

Availability of electricity and radio in Household

Although ownership of electricity and or/radio facilitate the acquisition of nutritional information allowing more successful allocation of resources to produce child health (Kandala, 2001), the availability of electricity and radio in households is not associated with the nutritional status of children in Egypt. However, only the availability of electricity seems to be mostly significant and has a positive effect on stunting, and underweight with separate analysis, and it seems to be significant on the LVM "nutritional status" in Nigeria. The reasons for these results may be, that mothers allocate their leisure time to radio or television, but it doesn't help improve the level of nutrition of their children. At same time, it reduces the length of time spent engaging in their children's affairs.

Antenatal visits and treatment during pregnancy

The variables access to health care, children of mothers who obtained clinical visits during pregnancy, had vaccines and treatment, have a positive and significant effect on malnutrition status. Therefore, health service investments are more effective in reducing stunting, wasting and underweight among indigenous communities. Our results indicate that children of mothers who had clinical visits and got a medical care during pregnancy are less likely to be stunted and to be underweight than their counterparts in Nigeria. On the other hand, in Egypt the covariate of *trepr* has a slight effect on the nutritional status, but *anvis* seems to be significant and has a positive influence on the underweight and wasting. The results with two indicators also indicate that the *anvis* has a positive effect in both countries.

In addition, the reason for the nonsignificant effects of toilet, mother's working status, and source of water in Egypt could be that, because most children are living in rural areas (66%) where many problems are found, such as the level of health or educational level of mothers, most women are not working (81%), and lack sanitation and water supply.

Child's age

In the analysis, it was discovered that the stunting of children increases gradually from 5-15 months of age in Egypt, where the minimum Z-scores of stunting is attained, then rises again through the reminder of the third year. In Nigeria, the situation of children who are stunted is quite similar; however, the deterioration in nutritional status is set between 5-20 months of age. Similarly, deterioration in child's *weight-for-height* sets in during first 4-5 months of age, as reported in much of the literatures, due to supplementation. However, it reaches its minimum level between age 13 and 15 months, then rises again and reaches its minimum level between age 26 and 28 months, which is later than the case of stunting in Egypt; and between age 16 and 18 months in Nigeria; which is earlier than the case of stunting. A sudden pick-up effect is noticeable from age 18 months until about 45 months, where it attains its maximum level in Nigeria. The pattern of underweight is similar to that of wasting in Nigeria. While in Egypt, deterioration in *weight-for-age* sets after 5 months of birth and increased dramatically until age 15 montes (which is the low stability level) and goes to be stable thereafter, however, in Nigeria an improvement commenced after age 20 or 25 months and rise gradually until age 50 months. Previous studies assumed that it is an average effect of low *height-for-age* and *weight-for-height* during this period of life (Adebayo, 2002).

The level of wasting suggests that insufficient food intake may be an important factor in the rise of malnutrition in both countries. In addition, the implication of this finding is that wasting is not clearly noticeable in the first four months of life. As soon as a child is fed with other supplementation such as liquids or other forms of diet which, due to the unhygienic source of preparation of such supplementations, may facilitate infections and diseases such as diarrhea, then acute malnutrition may set in. In other words, the introduction of liquids, such as water, sugar water, juice, tea, powdered or fresh milk, formula, and soiled food, takes place far earlier than the recommended age of about 6 months. This practice has a deleterious effect on nutritional status for many reasons. First, the liquid and solid foods offered are nutritionally inferior to breast milk. Second, the consumption of liquids and solid food decreases the infant's intake of breast milk, which in turn, reduces the mother's supply of milk (Breast milk production is determined, in part, by the frequency and intensity of suckling.) Third, feeding young infants liquids and solid food increases their exposure to pathogens and thus puts them at greater risk of diseases such as diarrhea (WHO 1994; Africa Nutrition Chartbooks, 1996).

Mother's BMI

A mother's nutritional status affects her ability to successfully carry, deliver, and care for her children and is of great concern in its own right. The analysis provides that virtually similar patterns are observed for all indices in both countries: approximately linear trends with positive slopes. Malnutrition in women is assessed using BMI. When the BMI of non-pregnant women falls below the suggested cut-off point, which is around $18.5kg/m^2$, malnutrition is indicated. Women who are malnourished (thinness or obesity) may have difficulty during childbirth and may deliver a child who can be wasted, stunted or underweight. The results indicate that there is an association between the thinness condition of the mother and the nutritional status of the child.

Mother's age at birth

The result show that the influence of mothers who are younger than 20 years is higher on the nutritional status of children in both countries.

Possible causes for this are due to childbirth among very young girls, whose bodies are not physically ready to endure the processes of childbirth. The problem is compounded by the fact that some African countries have poor obstetric care. Furthermore, these mothers could not reach health facilities, or, when they do, it is too late. Effective ways must be devised to delay age at first marriage and first birth. These two factors will almost certainly determine the number of children she will have in her lifetime. While early age at first birth has health implications, it also has economic implications. In addition, one study obtained in Nigeria reported that younger mothers (teenagers) are less likely in comparison to older mothers to breastfeed their children after birth, which means that the age of the mother at birth of a child influences whether the child will receive colostrum or not, which might affect the nutritional status of children (Adebayo, 2004). In other words, younger mothers are likely to positively affect their children's nutritional status. Moreover, other previous studies which were obtained in some developing countries have shown that some African countries do not allow girls back to go back school after they give birth. As a consequence, a girl who drops out of school will continue the cycle of poverty (Alderman et. al. 1997; Toroitich-Ruto, 1998).

Malnutrition among children by region

Prevalence of stunting, wasting and underweight among children by region in Egypt

The results indicate that the rural areas in the Nile delta and some other provenances there or in Lower Egypt are associated with malnutrition in children. One reason that, as some previous studies reported is that obesity among adults, particularly women, has reached very high proportions in Egypt in the last few years, while malnutrition rates in children (in the first two years of life) remain stubbornly high. The 1998 national food consumption survey reported underweight in 16.7% of 2-to 6-year-old children. Overweight and obesity affected 1.6% of 2- to 6-year old children. The prevalence of stunting in pre-school children ranged from 13% in Lower Egypt to 24% in Upper Egypt. At the same time, rates of early childhood malnutrition remain stubbornly stable and relatively high. The double burden of obesity and malnutrition is clearly evident. In addition, public awareness of the increasing prevalence of obesity and of diet-related chronic disease is increasing, and attention has turned to documenting the problem. On the other hand, most studies relating diarrhea and malnutrition have been conducted in economically marginal regions, where young children have high rates of diarrhea diseases and severely faltering growth. One study was conducted in an agricultural rural community in the Nile Delta. The population, although relatively uneducated, lived well. The villagers frequently owned their small homes or apartments, had access to municipal water, and often had modest luxuries such as radios and televisions. The incidence of diarrhea in children less than 3 years of age was moderate compared with that in other developing countries, and chronic diarrhea was uncommonly reported.

Prevalence of stunting, wasting and underweight among children by region in Nigeria

As reported in the 2003 NDHS, the trend in the nutritional status of Nigerian children has worsened with regard to stunting and wasting (from 36% in 1990 to 46% in 1999 for stunting and 11% in 1990 to 12% in 1999 for wasting). The results indicate that, mostly districts in the northeast and southeast positively associated with *height-for-age* and *weight-for-age*, while the districts in northwest are associated with *weight-for-height*. Providing a more complete picture, and based on our analysis, which reports above. The result also revealed striking regional variations, with the northeast, south and southeast in much worse situations in terms of stunting and underweight than the northwest and southwest. On the other hand, the children who live in the northwest part of the country are more likely to be wasted than their counterparts in other parts of country. These regional and zonal disparities may reflect the contribution of other factors, such as socio-cultural conditions and morbidity of children, as seen in chapter 5, in determining the nutritional status of children under age five. The high prevalence of stunting observed in the 2003 NDHS survey is in the context of large-scale deepening poverty and household food insecurity.[1] Severe rural poverty appears to be found in the southwest of Nigeria, in the north-center, and in the extreme northeast. These results are consistent with some previous studies which discuss the relation between poverty and malnutrition as persistent problems in Nigeria.

Summary and Conclusions

This study addresses the status of malnutrition in children under five in Egypt and Nigeria, using stunting, underweight and wasting as malnutrition indices. At the same time, it addresses the effects of different roles played by the various socioeconomic factors, such as mother's education, mother who is currently working, etc., in improving the children's nutritional condition. According to the results of this analysis, using separate geoadditive models and geoadditive latent variable models, the mother's education, sex of child, antenatal visits during pregnancy, source of water on undernutrition of child were important in both countries.

In addition, results showed that the place of residence, mother' working, type of toilet and availability of electricity and radio in household have negligible effects on the undernutrition of children.

We find that the methods identify the association of child's age, mother's age at birth and mother's BMI. It is found the children are at high risk during the first 15-20 months of life and then stabilize in Egypt, but the risk rises again between ages 25-50 months in Nigeria. The effect of BMI on the child's nutritional status is approximately linear with positive slope, which means that there is an association between the thinness condition of mothers and nutritional status. According to the mother's age at birth, it shows that younger mothers are less likely to affect their children's nutritional status positively.

It is found that children living in some provinces in the Nile Delta and Upper Egypt are having undernutrition problems. For Nigeria, the southeast regions and some regions in the southern part of the country are associated with undernutrition.

[1] Chronic Undernutrition among Children as an Indicator of Poverty, 2002.

Furthermore, the results using two indicators are quit similar to the results with three indicators in both countries.

Policy Implications

Integrate nutrition and family planning. This can help by delaying birth to ensure optimal growth (physical, psychological and emotional), and to raise the education level and therefore improve the socioeconomic status of mother and children.

Give more attention to some areas which have high rates of poverty, such as the Nile Delta, Upper Egypt and southeastern in Egypt, and some regions in the southern part of Nigeria. These areas are more likely to have a higher proportion of undernutrition compared to other areas, due to poor health facilities and complications during childbirth or even careless and misdiagnosis during hospital care. Therefore, the most important issues to address in these areas are health care, proper food, and raising the educational level of parents. Governments should improve socioeconomic conditions. Because, if living standards are improved, there will be better health care and a reduction in infant and, child diseases, child malnutrition and child mortality.

Do more research on ways to improve the nutritional status of households in these countries using indigenous inexpensive foods that are locally available. There is still a need for research and studies about nutrition and the important components of healthy eating to avoid the increase of illness caused by poor eating habits.

Parametrer	Mean	Std	2.5%	97.5%
		Factor Loadings		
stunting λ_{11}	0.655*	0.014	0.626	0.684
underweight λ_{21}	0.968*	0.008	0.952	0.983
		Parametric Indirect Effects		
male	-0.152*	0.027	-0.201	-0.099
anvis	0.077*	0.030	0.016	0.140
work	0.0137	0.036	-0.063	0.079
trepr	0.032	0.034	-0.036	0.098
elect	-0.047	0.045	-0.131	0.039
radio	0.008	0.133	-0.223	0.303
		Parametric Direct Effects		
water(a_{11})	-0.012	0.051	-0.117	0.078
educ(a_{12})	0.055*	0.022	0.011	0.104
toilet(a_{13})	-0.045	0.084	-0.212	0.104
urban(a_{14})	0.048	0.036	-0.020	0.121
water(a_{21})	-0.020	0.0413	-0.102	0.050
educ (a_{22})	0.043*	0.0148	0.005	0.063
toilet(a_{23})	-0.138*	0.067	-0.260	-0.0518
urban(a_{24})	0.023	0.0275	-0.026	0.068
		Smoothing Parameters		
Chage	0.003*	0.0071	0.0004	0.017
BMI	0.007*	0.009	0.0006	0.032
Mageb	0.002*	0.005	0.0003	0.013
reg	0.558*	0.241	0.244	1.177

Table 6.13: Estimates of factor loadings of the LVM3 with only two indicators in Egypt.

Parametrer	Mean	Std	2.5%	97.5%
		Factor Loadings		
stunting λ_{11}	1.147*	0.028	1.097	1.203
underweight λ_{31}	0.987*	0.0274	0.934	1.040
		Parametric Indirect Effects		
urban	0.0357	0.060	-0.357	0.152
anvis	0.346*	0.075	0.205	0.492
toilet	0.156	0.082	-0.013	0.313
elect	0.153*	0.058	0.033	0.269
		Parametric Direct Effects		
male(a_{11})	-0.242*	0.059	-0.357	-0.1372
work(a_{12})	0.087	0.064	-0.028	0.211
trepr(a_{13})	0.124	0.083	-0.044	0.290
water(a_{14})	0.065	0.086	-0.1033	0.241
educ(a_{15})	0.184*	0.067	0.055	0.330
radio(a_{16})	0.019	0.0365	-0.049	0.088
male(a_{21})	-0.057	0.045	-0.150	0.026
work(a_{22})	0.118*	0.053	0.0155	0.224
trepr(a_{23})	0.022	0.060	-0.090	0.137
water(a_{24})	0.0079	0.069	-0.124	0.139
educ (a_{25})	0.046	0.0529	-0.051	0.154
radio(a_{26})	0.028	0.029	-0.027	0.089
		Smoothing Parameters		
Chage	0.016*	0.018	0.064	0.143
BMI	0.004*	0.011	0.075	0.319
Mageb	0.135*	0.085	0.0003	0.009
reg	0.159*	0.054	0.081	0.291

Table 6.14: Estimates of factor loadings of the LVM3 with only two indicators in Nigeria.

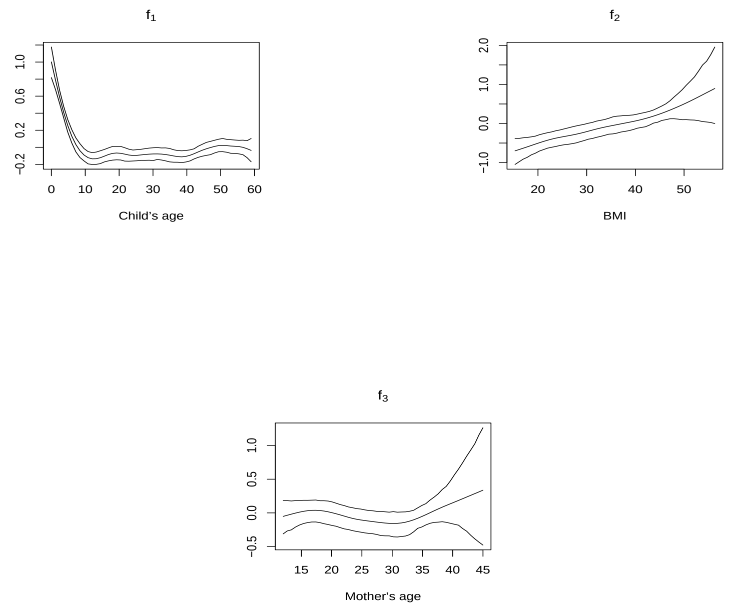


Figure 6.9: Non-linear effects from top to bottom: child's age, mother's BMI and mother's age at birth using only two indicators of a latent variable "Malnutrition status" of children for Egypt, using Bayesian latent variable model.

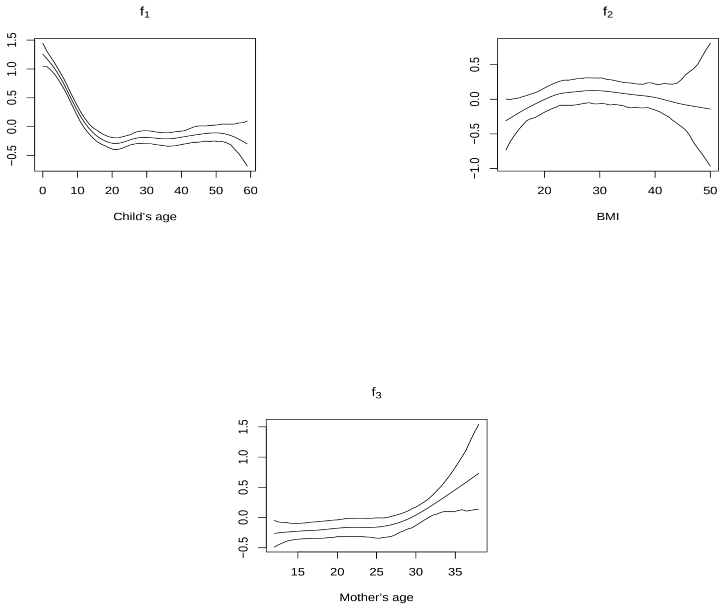


Figure 6.10: Non-linear effects from top to bottom: child's age, mother's BMI and mother's age at birth using only two indicators of a latent variable "Malnutrition status" of children for Nigeria, using Bayesian latent variable model for continuous responses.

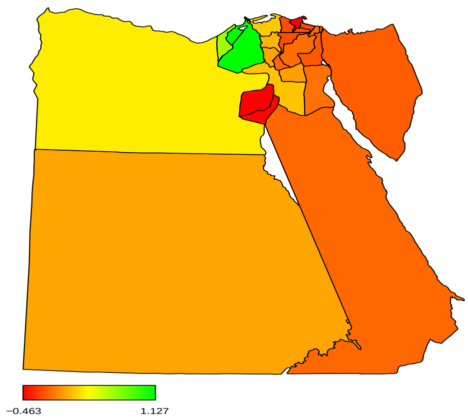


Figure 6.11: Posterior mean for latent variable model, using only two indicators of a latent variable "Malnutrition status" for Egypt.

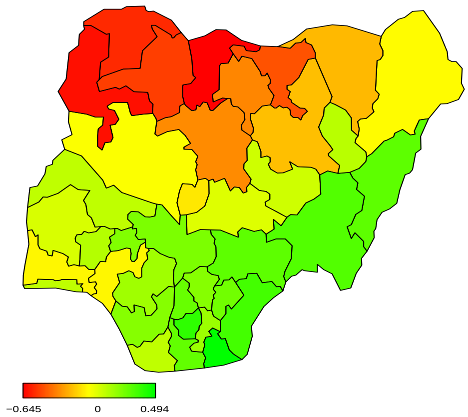


Figure 6.12: Posterior mean for latent variable model, using only two indicators of a latent variable "Malnutrition status" for Nigeria.

Chapter 7

Geoadditive Latent Variable Models for Disease and Nutrition Indicators

Abstract

In this chapter, several models with one or with two latent variables have been estimated using mixed indicators (the binary indicators "health status" and the continuous indicators "nutritional status") and a selection of covariates, that have been included in the previous analyses for childhood diseases and childhood undernutrition. The relationship between the indicators of diseases and the indicators of undernutrition based on the previous analyses (chapter 5 and chapter 6) is studied in this chapter.

7.1 Introduction

Based on the previous analyses of childhood diseases and childhood undernutrition, we started the modeling and the estimation using the binary indicators (*fever, diarrhea, and cough*) and the continuous indicators (*stunting, wasting and underweight*), with one latent or with two latent variables.

However, the results were not reasonable and some estimators did not convergent. In order to overcome this problem, we dropped the indicator of wasting which is strongly correlated with stunting and underweight, checked the results for five indicators again and obtained reasonable results. The first section discusses the latent variable models using one or two factors. In the second section, one latent variable with five indictors, including indirect parametric effects and direct parametric effects, is estimated. The third section discusses the estimation of two latent variables, using the same indicators and same covariates.

7.2 Latent Variable Models for Mixed Response Variables

In this application, we first introduce the scalar latent variable υ, "health and undernutrition status". We consider only a one-dimensional latent variable with different types of covariates (which were included in the previous chapters) in section 3. Extension to two-dimensional latent variables with the two types of responses and different types of covariates are presented in section 4. The latent variable for binary and metrical responses y_j, $j = 1, .., 5$. The response variables consist of five indicators: *fever, diarrhea, cough, stunting and wasting.*

Firstly, the *measurement model* using one latent variable is given by

$$y^*_{ij} = \lambda_0 + a'_j w_i + \lambda_j \upsilon_i + \epsilon_{ij}, \quad i = 1, .., n, \quad \epsilon_{ij} \sim N(0, \sigma^2_j) \tag{7.1}$$

for two metrical indicators and

$$y^*_{ij} = \lambda_0 + a'_j w_i + \lambda_j \upsilon_i + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, 1)$$

for the underlying variables y^*_{ij} corresponding to the three binary indicators y_{ij}, $j = 1, 2, 3$.

The form of the *structural model* is

$$\upsilon_i = u_i'\alpha + f_1(x_{i1}) + ... + f_3(x_{i3}) + f_{geo}(reg_i) + \delta_i \tag{7.2}$$

The models include the direct vector of covariates w_i for each individual response variable. The direct vector w_i includes the categorical covariates *water, educ, toilet, urban, trep and elect* in the LVM for Egypt; but in the case of Nigeria, it includes the covariates *male, educ, radio, and water.* The indirect vector u includes *male, anvis, work, and radio* in the latent variable mixed models for Egypt, and *urban, work, terp, avis, toilet, and elect* for Nigeria.

Secondly, the *measurement model* using two latent variables is given by

$$\begin{pmatrix} y_{i1}^* \\ y_{i2}^* \\ y_{i3}^* \\ y_{i4}^* \\ y_{i5}^* \end{pmatrix} = \begin{pmatrix} \lambda_{i1} \\ \lambda_{i2} \\ \lambda_{i3} \\ \lambda_{i4} \\ \lambda_{i5} \end{pmatrix} + \begin{pmatrix} a_1' \\ a_2' \\ a_3' \\ a_4' \\ a_5' \end{pmatrix} . \begin{pmatrix} w_1 & w_2 & w_3 & w_4 & w_5 \end{pmatrix} + \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \\ \lambda_{31} & \lambda_{32} \\ \lambda_{41} & \lambda_{42} \\ \lambda_{51} & \lambda_{52} \end{pmatrix} . \begin{pmatrix} \upsilon_1 \\ \upsilon_2 \end{pmatrix} + \begin{pmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \epsilon_{i3} \\ \epsilon_{i4} \\ \epsilon_{i5} \end{pmatrix} \tag{7.3}$$

The *structural model* for the analysis uses two latent factors:

$$\begin{pmatrix} \upsilon_1 \\ \upsilon_2 \end{pmatrix} = \begin{pmatrix} u'_{i1}\alpha_1 \\ u'_{i2}\alpha_2 \end{pmatrix} + \begin{pmatrix} f_{11}(Chage_i) \\ f_{21}(Chage_i) \end{pmatrix} + \begin{pmatrix} f_{12}(BMI_i) \\ f_{22}(BMI_i) \end{pmatrix} + \begin{pmatrix} f_{13}(Mageb_i) \\ f_{23}(Mageb_i) \end{pmatrix} \quad (7.4)$$

$$+ \begin{pmatrix} f_{14}(reg_i) \\ f_{24}(reg_i) \end{pmatrix} + \begin{pmatrix} \delta_{i1} \\ \delta_{i2} \end{pmatrix}$$

In this application, the following five response variables are included as indicators for childhood diseases and undernutrition.

Response variables

fever: 1 if child had disease 2 weeks prior to the survey and 0 otherwise (reference category).

diarrhea: 1/0 (reference).

cough: 1/0(reference).

stunting: Height-for-age which indicates stunting.

underweight: Weight-for-age an indication of underweight.

The following covariates were considered in the analysis in both countries:

Metrical covariates

Chage: Child's age in months.

BMI: Mother's body mass index.

Mageb: Mother's age at birth.

Categorical covariates (in effect coding)

male: Child's sex: male or female (reference category).

educ: Mother's educational attainment: incomplete primary, complete primary, and incomplete secondary school or complete secondary school and higher eduction (reference category).

trepr: Whether mother had treatment during pregnancy: yes or no (reference category).

anvis: Whether mother had antenatal care: yes or no (reference).

water: Source of drinking water: controlled water or no (reference category).

toilet: Has flush toilet at household: yes or no (reference category).

urban: Location where respondent lives: urban or rural (reference category).

radio: Has a radio at household: yes or no (reference category).

elect: Has electricity: yes or no (reference category).

work: Mother's current working status: working or not (reference).

Spatial covariate

reg: Governorate or region where respondent resides.

7.3 Model Estimation with One Factor Analysis

In this section, we investigate how the indicators of diseases and undernutrition can be interpreted as indicators of the latent variables "health status" and "nutritional status" of children, respectively and how much of the variation of latent variable can be explained through the predictors. On the other hand, this concept does not only allow us to analyze the impact of covariates on the indicators of health and nutritional status, but also allows us to know the correlation among the indicators. In order to decide which of the covariates should be included in the measurement model as direct parametric covariates, or in the structural equation as indirect effect via their impact on the latent variable, the results of LVM3 and LVM2 in chapter 5 and chapter 6, respectively are taken into account. In other words, as mentioned in chapter 5 and chapter 6, the covariates *male, antenatal visit, radio* and *work* were associated with childhood diseases and childhood undernutrition, so we kept them in the structural equation in the analysis of Egypt. The same is done for Nigeria. Based on the previous results of chapter 5 and chapter 6 the covariates *urban, work, antenatal visit, treatment during pregnancy, toilet*, and *electricity* are included in the structural model.

Our analysis started using only one latent variable. The results for the estimation of factor loadings, parametric indirect and direct effects for Egypt and Nigeria are presented in tables 7.1 and 7.2.

The factor loadings (see table 7.1) show that the latent variable has a stronger influence on the first three indicators that belong to the health status than on the nutritional status (*stunting* and *underweight*). The parametric indirect effects for *male, antenatal visit and work* have a significant effect on the latent variable. Regarding the parametric direct effects, covariate *urban* is associated with indicators 2 *(cough)*, 3 *diarrhea*, 4 *(stunting)* and 5 (underweight), whilst *treatment during pregnancy* is associated with the second indicator; and the *education level of mother* has a positive effect on the indicators of *stunting* and *underweight.*

In addition, none of the covariates which have parametric effects were associated with the indicators of *fever.* For Nigeria, the results (table 7.2) indicate the same conclusion as for Egypt, which assumed that the estimates of factor loadings for the diseases affect the latent variable more than the indicators of undernutrition. The results show that the indicators of undernutrition have a slight significant effect on the latent variable. The results of the indirect parametric covariates show that only *urban* and *treatment during pregnancy* have significant effect on the latent variable. As for the direct parametric covariates, *male, education level* and *radio* are associated with the indicator 4 *(stunting)*, whilst only *the level of education* is associated with the indicator 2 *(cough).*

With regards to the nonparametric effects, figures 7.1 and 7.4 show the nonlinear and spatial effects for Egypt and Nigeria, respectively. The patterns of the nonlinear effects are very similar to the analysis of diseases (see figure 5.13 and figure 5.14, chapter 5). These results are reasonable and expected, because the indicators of diseases are clearly represented through the latent variable, so the results are consistent with the results of the childhood diseases. The nonlinear effect of child's age indicates that the prevalence of diseases was found to be highest among children 0-12 months of age. As for the effect of a mother's BMI, it has a slight effect on the latent variable; however, there is a higher effect through the interval between 27-30. The pattern of mother's age shows that younger mothers have a higher effect on the latent variable than their counterparts. There is no new suggestion found in the spatial effect for Egypt (second right panel of figure 7.1). It indicates that higher risks are associated with some rural areas in the Nile Delta and in Sinai as well. The results are consistent with the results of childhood disease for Egypt (see chapter 5, section 4).

The patterns of the nonparametric effects for the nonlinear effects of a child's age, mother's BMI and a mother's age as well as child's residence on the latent variable health status and nutritional status are quite similar to the patterns of the analysis for childhood diseases in Nigeria (see chapter 5, section 4). In particular, the health status of children worsens until about 12 months of age. The effect of BMI seem to be a little higher for mothers with a BMI under 20 and after that, it is slight. Children from younger mothers, as in Egypt, are more likely to have problems, particularly in their health status. We have found that the high risk of the latent variable health and nutritional status is associated with the northeastern part of Nigeria, as already mentioned and indicated in the previous analyses of childhood diseases (see chapter 5, section 4).

7.4 Model Estimation with Two Latent Variables

In this section, we analyze determinants of childhood diseases and childhood undernutrition using two latent variables instead of one.

The factor loadings estimates (tables 7.3 and 7.4) show that the first latent variable loads onto the first three indicators, whilst indicators 4 and 5 are explained by the second latent variable. This was to be expected, because the two different sets of indicators are supposed to measure two different latent constructs. Both factor loadings and coefficients of the parametric indirect covariates of the first latent factor are very similar to the estimates of the single latent factor model given in 7.1. Regarding the factor loadings of the second latent variable, the indicator *underweight* is deserving of notice due to its high factor loading of 0.975 in Egypt; but in Nigeria, the second latent variable has a highest influence on the indicator *stunting*, with factor loading of 1.16. Results for Egypt show that the influence of the covariates *anvis*, *male*, *radio* and *work* is noticeable for the first latent

Parametrer	Mean	Std	2.5%	97.5%
		Factor Loadings		
1. Fever λ_{11}	1.247*	0.089	1.094	1.420
2. Cough λ_{21}	0.811*	0.047	0.724	0.901
3. Diarrhea λ_{31}	0.816*	0.043	0.734	0.897
4. stunting λ_{41}	−0.132*	-0.134	-0.182	-0.084
5. underweight λ_{51}	−0.133*	0.021	-0.015	-0.074
		Parametric Indirect Effects		
male	0.168*	0.038	0.036	0.247
anvis	0.221*	0.064	0.098	0.339
work	0.123*	0.053	0.0168	0.238
radio	-0.164	0.108	-0.275	0.072
		Parametric Direct Effects		
water(a_{11})	0.122	0.088	-0.037	0.295
educ(a_{12})	-0.065	0.049	-0.157	0.028
toilet(a_{13})	-0.107	0.128	-0.452	0.116
urban(a_{14})	0.047	0.068	-0.082	0.183
trepr(a_{15})	0.076	0.097	-0.108	0.272
elect(a_{16})	-0.313	0.279	-0.838	0.211
water(a_{21})	0.061	0.067	-0.065	0.195
educ(a_{22})	-0.055	0.041	-0.140	0.015
toilet(a_{23})	-0.089	0.120	-0.348	0.108
urban (a_{24})	−0.22*	0.063	-0.348	-0.098
trepr(a_{25})	0.193*	0.073	0.039	0.340
elect(a_{26})	0.191	0.071	-0.467	0.532
water(a_{31})	-0.02	0.067	-0.178	0.112
educ (a_{32})	-0.033	0.038	-0.128	0.037
toilet(a_{33})	-0.029	0.116	-0.277	0.184
urban(a_{34})	0.151*	0.056	0.044	0.265
terpr(a_{35})	0.015	0.079	-0.139	0.165
elect(a_{36})	-0.096	0.231	-0.516	0.357
water(a_{41})	0.006	0.050	-0.090	0.108
educ(a_{42})	0.066*	0.026	0.015	0.118
toilet(a_{43})	0.029	0.084	-0.121	0.208
urban (a_{44})	0.123*	0.037	0.054	0.203
terpr(a_{45})	-0.009	0.057	-0.108	0.098
elect(a_{46})	-0.171	0.168	-0.519	0.153
water(a_{51})	0.013	0.039	-0.065	0.099
educ (a_{52})	0.052*	0.022	0.013	0.095
toilet(a_{53})	-0.035	0.070	-0.197	0.119
urban (a_{54})	0.13*	0.036	0.069	0.199
trepr (a_{55})	-0.023	0.0335	-0.126	0.079
elect(a_{56})	-0.088	0.1455	-0.351	0.200

Table 7.1: Estimates of factor loadings, parametric indirect and direct effects of the LVMM with one latent variable for Egypt.(*: Statistically significant at 2.5%).

Parametrer	Mean	Std	2.5%	97.5%
		Factor Loadings		
1. Fever λ_{11}	0.821^*	0.081	0.682	0.989
2. Cough λ_{21}	0.651^*	0.063	0.538	0.781
3. Diarrhea λ_{31}	0.896^*	0.084	0.741	1.087
4.stunting λ_{41}	-0.262^*	0.046	-0.348	-0.171
6. underweight λ_{51}	-0.21^*	0.028	-0.246	-0.136
		Parametric Indirect Effects		
urban	-0.179^*	0.079	-0.326	-0.017
work	0.004	0.070	-0.126	0.147
trepr	0.204^*	0.074	0.053	0.331
anvis	-0.039	0.085	-0.204	0.126
toilet	-0.111	0.100	-0.325	0.078
elect	-0.018	0.077	-0.171	0.127
		Parametric Direct Effects		
male(a_{11})	-0.006	0.068	-0.141	0.128
educ (a_{12})	0.024	0.077	-0.125	0.181
radio(a_{13})	-0.030	0.040	-0.105	0.047
water(a_{14})	0.044	0.095	-0.130	0.243
male(a_{21})	0.016	0.061	-0.102	0.134
educ (a_{22})	0.151^*	0.070	0.020	0.282
radio(a_{23})	0.007	0.039	-0.069	0.086
water(a_{24})	-0.059	0.093	-0.240	0.113
male(a_{31})	0.128	0.078	-0.030	0.274
educ (a_{32})	-0.118	0.094	-0.318	0.051
radio(a_{33})	-0.044	0.044	-0.136	0.041
water(a_{34})	-0.050	0.111	-0.249	0.183
male(a_{41})	-0.218^*	0.067	-0.347	-0.100
educ (a_{42})	0.449^*	0.0718	0.308	0.584
radio(a_{43})	0.090^*	0.040	0.012	0.166
water(a_{44})	0.093	0.095	-0.090	0.274
male(a_{51})	0.063	0.052	-0.032	0.165
educ (a_{52})	-0.016	0.056	-0.122	0.085
radio(a_{53})	0.018	0.030	-0.037	0.076
water(a_{54})	-0.048	0.069	-0.193	0.087

Table 7.2: Estimates of factor loadings, parametric indirect and direct effects of the LVMM with one latent variable for Nigeria.

variable, whilst the second latent variable is associated with *anvis* and *male*. The results of the parametric direct covariates are quite similar to the estimates with a single latent variable. The results are also consistent with the analysis of childhood diseases and childhood malnutrition (see ch5 and ch6). For Nigeria, the influence of *urban* and *trepr* is associated with the first latent variable, however, the second latent variable is more influenced by the covariates of *avis* and *elect*; whilst the coefficients of the parametric direct covariates are very slight and nonsignificant, with the exception of the *child's sex* which has a significant effect on the indicator *stunting* as the covariate of *education level* which has a significant effect on indicators of *diarrhea* and *stunting*. Furthermore, the results provided in table 7.4 are very similar to the estimates of the single latent variable in 7.1. The results of Nigeria using two latent factor are also consistent with the earlier results which were obtained in chapter 5 and chapter 6. The patterns of the covariates child's age, mother's BMI and mother's age resemble the patterns of the model with one latent variable (left panels of figure 7.2 and figure 7.5) drawn in figures 7.1 and 7.4, whilst the influence of these covariates on the second latent variable looks different. Apparently, the nonlinear effects on the second latent variable are associated with the indicators of nutritional status.

A finding is that, the nutritional status of children worsens after around 5 month of birth until about 25 months of age in Egypt and until about 20 months of age in Nigeria, whilst the effect of BMI is very slight in both countries. The effect of young mothers on the nutritional status of children is higher compared to that of older mothers (see right panel of figure 7.2). However, in Nigeria the effect of mother's age on the nutritional status of children is slight. The estimates of the spatial effect for the first latent variable (left panels of figures 7.3 and 7.6) are similar but slightly different compared to the estimates for the one latent variable model in the bottom right panel of figures 7.1 and 7.4, respectively. The reason for the slightly different parameter estimates and their significance lies in the fact that the number of observations in the two-factor model is different from the number of observation in one latent variable. The spatial effect of the second latent variable (left panels of figures 7.3 and 7.6) is associated with the indicators of nutritional status and it depicts that the effect of some rural provinces in the Nile Delta on the nutritional status of children in Egypt is high. As for Nigeria, the pattern of the spatial effect indicates that the southeastern part of country is associated with a significant effect on the nutritional status of children.

This section demonstrates that the model we used can estimate the influence of covariates on more than one latent variable. On the other hand, the results of the model with two latent variables are consistent with the results of the single latent variable. Furthermore, the results of the latent variable models using the mixed indicators (health status indicators and the nutritional status indicators) are consistent with the earlier results of the childhood diseases and malnutrition which were obtained in ch5 and ch6, respectively.

Parametrer	Mean	Std	2.5%	97.5%
Factor Loadings of First LV				
1.Fever λ_{11}	1.221*	0.092	1.07	1.496
2.Cough λ_{21}	0.810*	0.0438	0.721	0.903
3.Diarrhea λ_{31}	0.816*	0.0441	0.737	0.914
4.stunting λ_{41}	-0.066*	0.0179	-0.106	-0.019
5.underweight λ_{51}	-0.048	0.0109	-0.051	0.039
Factor Loadings of Second LV				
1.Fever λ_{12}	0.000	0.000	0.000	0.000
2.Cough λ_{22}	0.033	0.023	-0.058	0.046
3.Diarrhea λ_{32}	-0.031	0.0222	-0.054	0.039
4.stunting λ_{42}	-0.657*	0.0145	-0.325	-0.260
5.underweightλ_{52}	-0.975*	0.007	1.061	1.099
Parametric Indirect Effects of First LV				
male	0.1319*	0.0385	0.056	0.207
anvis	0.206*	0.0428	0.127	0.288
work	0.108*	0.0510	0.0133	0.205
radio	-0.737*	0.366	-1.359	-0.0512
Parametric Indirect Effects of Second LV				
male	0.152*	0.025	0.101	0.200
anvis	-0.085*	0.0296	-0.140	-0.025
work	-0.0159	0.036	-0.0870	0.052
trepr	-0.020	0.036	-0.089	0.054
elect	0.040	0.047	-0.0527	0.123
radio	0.147	0.133	-0.0982	0.398
Parametric Direct Effects of Both LVs				
water(a_{11})	0.138	0.085	-0.0257	0.291
educ (a_{12})	-0.051	0.0504	-0.149	0.038
toilet(a_{13})	-0.110	0.145	-0.398	0.175
urban(a_{14})	0.031	0.068	-0.1002	0.164
trepr(a_{15})	0.082	0.0917	-0.095	0.262
elect(a_{16})	-0.311	0.267	-0.839	0.208
water(a_{21})	0.079	0.073	-0.064	0.221
educ(a_{22})	-0.053	0.0413	-0.130	0.024
toilet(a_{23})	-0.059	0.125	-0.297	0.176
urban (a_{24})	-0.211*	0.0610	-0.338	-0.090
trepr(a_{25})	0.192*	0.077	0.044	0.349
elect(a_{26})	0.018	0.251	-0.512	0.464
water(a_{31})	-0.023	0.074	-0.162	0.119
educ (a_{32})	-0.036	0.040	-0.113	0.028
toilet(a_{33})	-0.023	0.119	-0.282	0.191
urban(a_{34})	0.152*	0.059	0.0446	0.276
trepr(a_{35})	0.014	0.08	-0.136	0.170
elect(a_{36})	-0.112	0.234	-0.595	0.360
water(a_{41})	-0.019	0.049	-0.109	0.084
educ(a_{42})	0.061*	0.0219	0.023	0.104
toilet(a_{43})	-0.007	0.083	-0.175	0.155
urban (a_{44})	0.036	0.0352	-0.033	0.099
water(a_{51})	-0.025	0.031	-0.0718	0.037
educ (a_{52})	0.048*	0.009	0.0278	0.064
toilet(a_{53})	-0.02	0.068	-0.194	0.132
urban(a_{54})	0.11*	0.034	0.065	0.194
trepr(a_{55})	-0.068	0.042	-0.151	0.013
elect(a_{56})	-0.73	0.132	-0.327	0.160

Table 7.3: Estimates of factor loadings of the LVMM with two latent variable and only five indicators for Egypt.

Parametrer	Mean	Std	2.5%	97.5%
Factor Loadings of First Latent Variable				
1. Fever λ_{11}	0.957*	0.083	0.821	1.146
2. Cough λ_{21}	1.032*	0.091	0.868	1.208
3. Diarrhea λ_{31}	0.77*	0.065	0.665	0.901
4.stuntingλ_{41}	-0.025	0.048	-0.118	0.073
5. underweight λ_{51}	-0.155*	0.037	-0.226	-0.090
Factor Loadings of Second Latent Variable				
1. Fever λ_{12}	0.000	0.000	0.000	0.000
2. Cough λ_{22}	0.253*	0.045	0.164	0.336
3. Diarrhea λ_{32}	-0.088*	0.0355	-0.1588	-0.014
4.stunting λ_{42}	1.165*	0.028	1.109	1.224
5.underweightλ_{52}	0.958*	0.0238	0.910	1.006
Parametric Indirect Effects of First LV				
urban	-0.144*	0.067	-0.277	-0.020
work	0.010	0.068	-0.108	0.160
trepr	0.243*	0.074	0.091	0.380
avis	-0.037	0.076	-0.171	0.111
toilet	-0.075	0.099	-0.287	0.109
elect	-0.023	0.074	-0.170	0.121
Parametric Indirect Effects of Second LV				
urban	0.021	0.060	-0.103	0.141
work	0.105	0.060	-0.014	0.219
trepr	0.093	0.0613	-0.030	0.218
anvis	0.359*	0.063	0.234	0.474
toilet	0.143	0.080	-0.012	0.304
elect	0.159*	0.064	0.029	0.287
Parametric Direct Effects of Both LV				
male(a_{11})	-0.0073	0.066	-0.135	0.1230
educ (a_{12})	-0.039	0.078	-0.186	0.130
radio(a_{13})	-0.052	0.042	-0.137	0.034
water(a_{14})	0.041	0.103	-0.168	0.245
male(a_{21})	0.032	0.074	-0.104	0.168
educ (a_{22})	0.068	0.087	-0.100	0.242
radio(a_{23})	-0.024	0.048	-0.115	0.068
water(a_{24})	-0.056	0.110	-0.261	0.185
male(a_{31})	0.115	0.0707	-0.023	0.256
educ (a_{32})	-0.154*	0.081	-0.312	-0.006
radio(a_{33})	-0.066	0.038	-0.142	0.011
water(a_{34})	-0.06	0.102	-0.266	0.137
male(a_{41})	-0.242*	0.063	-0.374	-0.119
educ (a_{42})	0.185*	0.066	0.0565	0.326
radio(a_{43})	0.028	0.039	-0.052	0.105
water(a_{44})	0.072	0.0897	-0.102	0.230
male(a_{51})	-0.061	0.047	-0.1472	0.042
educ (a_{52})	0.048	0.054	-0.052	0.153
radio(a_{53})	0.027	0.029	-0.030	0.082
water(a_{54})	-0.005	0.072	-0.144	0.139

Table 7.4: Results of LVMM using 2 latent variable for Nigeria.

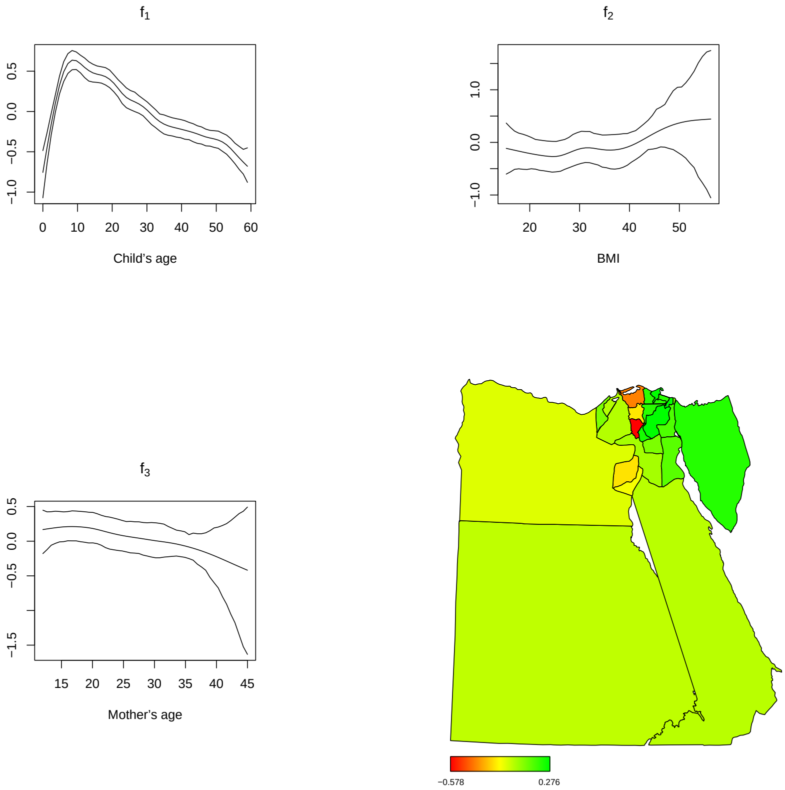


Figure 7.1: Non-linear effects from top to bottom: child's age, mother's BMI, mother's age at birth and spatial effects (for model LVMM using 5 indicators), on the indicators of a latent variable "health status" and "undernutrition status" of children for Egypt using only one latent variable.

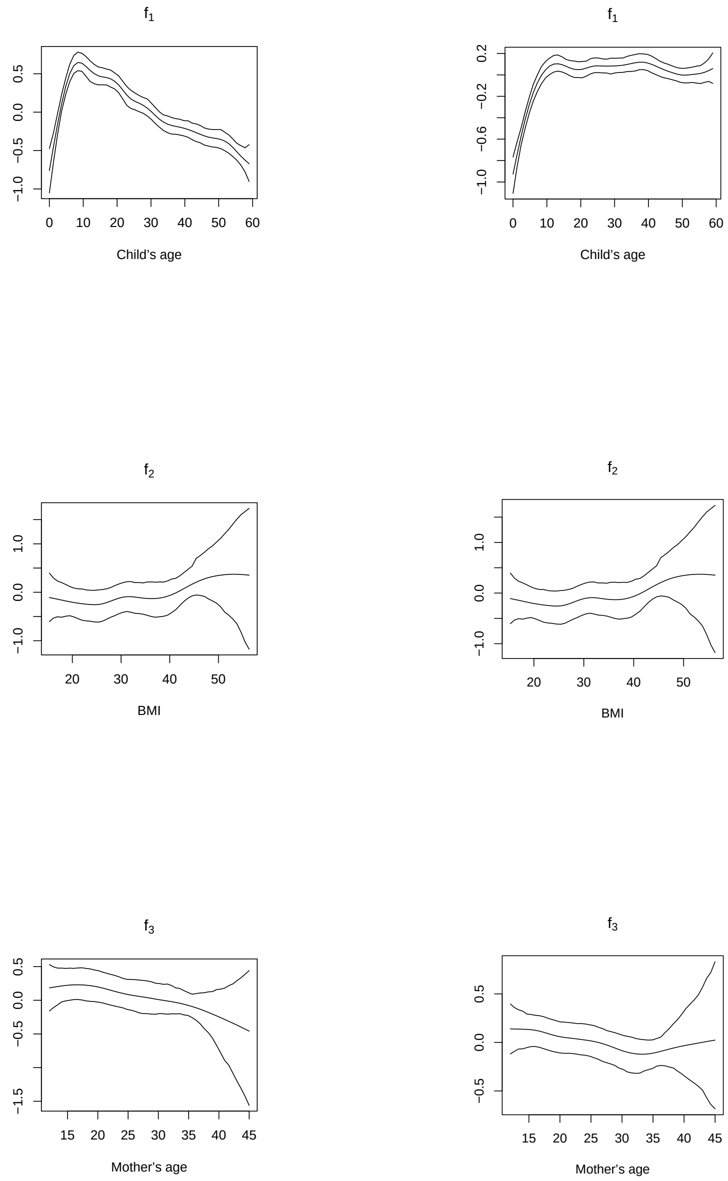


Figure 7.2: Estimates of nonparametric effect of nonlinear covariates from top to bottom: child's age, mother's BMI, and mother's age at birth for the first (left) and second latent variable for Egypt.

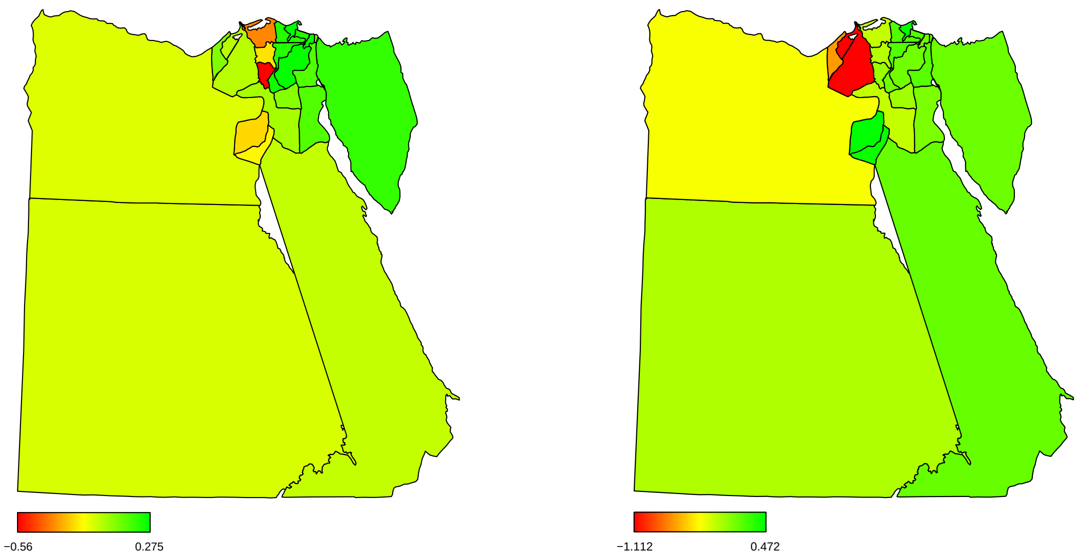


Figure 7.3: Estimates of the nonparametric effect of spatial covariate for the first (left) and second (right) for Egypt.

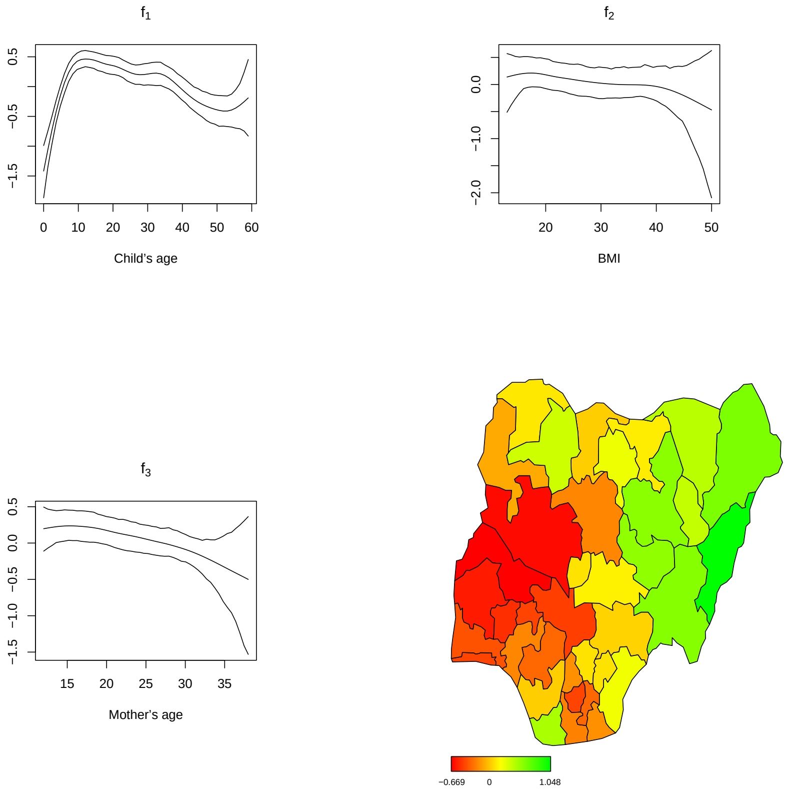


Figure 7.4: Non-linear effects from top to bottom: child's age, mother's BMI, mother's age at birth and spatial effects (for model LVMM using 5 indicators), on the indicators of a latent variable "health status" and "undernutrition status" of children for Nigeria using only one latent variable.

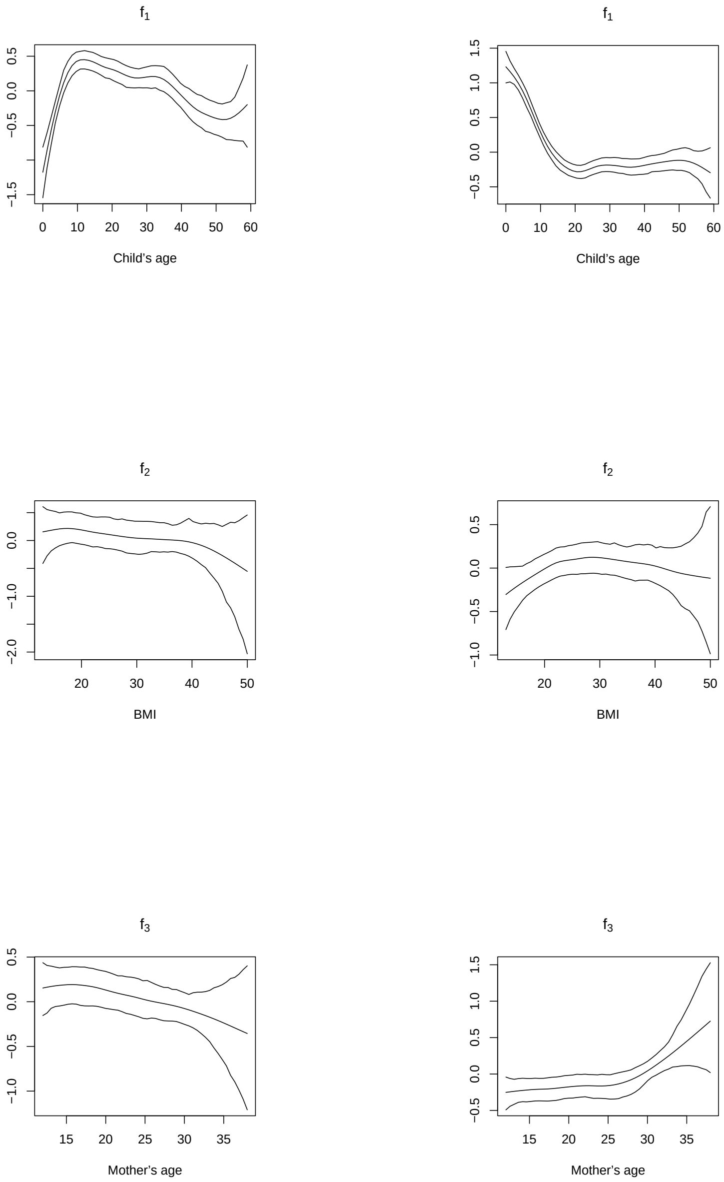


Figure 7.5: Estimates of nonparametric effect of nonlinear covariates from top to bottom: child's age, mother's BMI, and mother's age at birth for the first (left) and second latent variable for Nigeria.

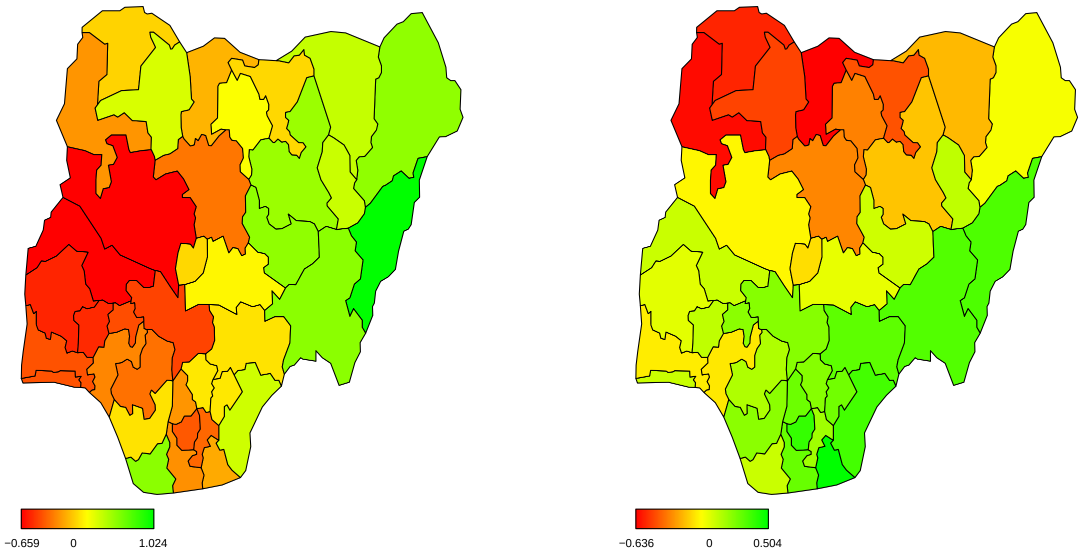


Figure 7.6: Estimates of the nonparametric effect of spatial covariate for the first (left) and second (right) for Nigeria.

Bibliography

Adebayo, S.B. (2004). Bayesian Geoadditive Modelling of Breastfeeding Initiation in Nigeria. Journal of Applied Econometrics 19: 267-281 (2004).

Adebayo S.B. (2002). Semiparametric Bayesian Regression for Multivariate Responses. Doctoral Thesis, University of Munich, Germany.

Adebayo S.B., and Fahrmeir, L. (2002). Analyzing Child Mortality in Nigeria with Geoadditive Survival Models. SFB 386, Discussion Paper 303, University of Munich, Germany.

Aguilar, O., West, M. (2000). Bayesian dynamic factor models and variance matrixdis- counting for portfolio allocation. Journal of Business and Economic Statistics 18, 338-357.

Alderman H, JR Behrman, V Lavy and R Menon. Child Nutrition, Child health and School Enrollment.: A longitudinal Analysis. Policy research Working Paper.1700. The World Bank. 1997. 1-6.

African Nutritition Chartbooks (1996). Nutrition of Infants and young Children in Mali. Macro international Inc. Calverton, Maryland, U.S.A.

Anderson, T.W., Rubin, H.(1956). Statistical inference in factor analysis.In Neyman, J. (Ed.), Proceedings of the Third Berkeley Symposiumon Mathematical Statistics and Probability 5, 111-150. Berkeley, University of California Press.

Ansari, A., Jedidi, K. (2000). Bayesian factor analysis formultilevel binary observations. Psychometrika 65, 475-496.

Arminger, G., Küsters, U. (1988). Latent trait models withind icators of mixed measure- ment level. In Langeheine, R., Rost, J.(Eds.) Latent Trait and Latent Class models, 51-73.Plenum, NewYork.

Arminger, G.and Muthen, B.O. (1998). A Bayesian approach to nonlinear latent variable models using the Gibbs sampler and the Metropolis-Hastings algorithm. Psychometrika 63, $271 - 300$.

Bartholomew, D.J. (1981). Posterior analysis of the factor model. British Journal of Math- ematical and Statistical Psychology 34, 93-99.

Bartholomew, D.J. (1987). Latent variable models and factor analysis. Charles Grif- fin and Company Ltd., London, Oxford University Press, New York.2nd edition: Bartholomew, D.J., Knott, M.(1999). Arnold, London.

Bateman OM., smith S. (1991). A comparison of the health effects of water supply and sanitation in urban and rural Guatemala. In: Proceeding of DHS world confrence, Washington, D.C. 5-7 August 1991, v. l Columbia, MD: Macro International InC. 1991: 745-56.

BayesX (2005). Software for Bayesian inference by Brezger, A., Kneib, T., Lang, S.

URL:http://www.stat.uni-muenchen.de/bayesx/bayesx.html.

Besag, J., York, Y. and Mollie, A. (1991). Bayesian Image Restoration with Two Applications in Spatial Statistics (with Discussion). Ann. Inst. Statist. Math., 43, 1-59.

Besag, J. and Kooperberg, C. (1995). On conditional and intrinsic autoregressions. Biometrika 82, 73-746.

Black, R.E. (1984). Diarrhea diseases and child morbidity and mortality, Population Development Revue suppl.(10):41-200.

Bharath Kummar, K. (2005). Child Undernutrition Morbidity and Poverty: Evidence on Cumulative Risk for Children in Tamil Nadu and Bihar. International Institute for population Sciences-Mumbai.

Boerma, J., R.E. Black, A.E. Sommerfely, S.O. Rustein, and G.T. Bicego (1991). Accuracy and Completeness of Mother's Recall of Diarrhea Occurrence in Preschool Children in Demographic and Health Surveys International Journal of Epidemiology, 20(4):1073-1080.

Bolstad M.W. and Manda O. S. (2001). Investigating Child Mortality in Malawi Using Family and Community Ramdon Effects: A Bayesian Analysis. Journal of American Statistical Associationt Vol. 96 No(453).

Borooah, V. (2002). The Role of Maternal Literacy in Reducing the Risk of Child Malnutrition in India. ICER - International Centre for Economic Research.

Bollen, K.A.(1989). Structural equations with latent variables. John Wiley, New York.

Boomsma, A. (1985). Non convergence, improper solutions, and starting values in LISREL maximum likelihood estimation. Psychometrika 50, 229-242.

Branca, F. and Ferrari, M (2002). Impact of Micronutrient Deficiencies on Growth: The Stunting Syndrome, National Institute for Food Nutrition Research.

Brezger, A.and Lang, S.(2005). Generalized structured additive regression based on Bayesian P-splines. Computational Statistics and Data Analysis 50, 967-991.

Caputo, A., R. Foraita, S. Klasen, and I. Pigeot. (2002). Undernutrition in Benin: An Analysis based on Graphical Models. Social Science and Medicine (forthcoming).

Christiaensen, L. and Alderman, H. 2001. Child Malnutrition in Ethiopia: Can Maternal Knowledge Augment the Role of Income, World Bank, 2001.

Chib,S. , Nardari, F.and Shephard, N.(2002). Markov chain Monte Carlo methods for stochastic volatility models, Journal of Econometrics, 108, $281-316$.

Costello, M.A., L.C. Lleno and E.R. Jensen. (1996). Determinants of major diseases and their Treatment in Philippines: Findings from the 1993 National Demographic Survey. Asian- Pacific Population Research Reports (9): August.

De Boor, C. (1978). A Practical Guide to Splines. Springer, Berlin.

Demographic and Health Survey (EDHS) for Egypt (2003).

Demographic and Health Survey (NDHS) for Nigeria (2003).

Denison, D. G. T., Mallick, B. K. and Smith, A. F. M. (1998). Automatic Bayesian curve fitting. Journal of the Royal Statistical Society B 60, 333-350.

Diggle, P., Moyeed, R., and Thomson, M. 2002. Childhood Malaria in the Gambia: A Case- Study in Model-Based Geo statistics. Applied Statistics 51: 4: 493-506.

Dierckx, P. (1993). Curve and Surface Fitting with Splines. Clarendon Press, Oxford.

Diggle, P. J., Ribeiro Jr., P. J. and Christensen O. F. (2003). An introduction to model based geostatistics. In: Møller, J. (ed.): Spatial Statistics and Computational Methods. Springer, New York.

El-Sayed , N.; Mohamed, A.G.; Nofal, L.; Mahfouz, A.; Abou Zeid, H. (2001) ICDDR,B: Centre for Health and Population Resear. J HEALTH POPUL NUTR, 2001 Dec;19(4): 275-280.

Eilers, P.H.C. and Marx, B.D.(1996). Flexible smoothing with B-splines and penalties (with comments). Statistical Science 11, 89-121.

Fahrmeir, L. and Lang, S. (2001a). Bayesian Inference for Generalized Additive Mixed Models Based on Markov Random Field Priors. Applied Statistics (JRSS C), 50, 201-220.

Fahrmeir, L., Lang, S. (2001b). Bayesian semiparametric regression analysis of multicate- gorical time-spacedata. Annals of the Institute of Statistical Mathematics 53, 10-30.

Fahrmeir, L., Kneib, T.and Lang, S. (2004). Penalized structured additive regression of space-time data:a Bayesian perspective. Statistica Sinica 14, 731-761.

Fahrmeir, L., and Raach, A. (2006). A Bayesian semiparametric latent variable model for mixed responses. Discussion paper 471, revised for Psychometrika.

Fahrmeir, L., Tutz, G. (2001). Multivariate statistical modelling based on generalized linear models. Springer, New York.

Folasade, I. and Adewale, O. (2000). Ecological and cultural barriers to treatment of childhood diarrhea in riverine areas of Ondo State, Nigeria. Social Science and Medicine, Volume 50, Number 7, 1 April 2000, pp. 953-964(12).

Filmer, D. and L. Pritchett. (2001). Estimating Wealth Effects Without Income of Expenditure Data-Or Tears: An Application To Educational Enrollment In States of India, Demography 38(1):115-132.

Folasade Iyun, B., Adewale Oke, E. (2000). Ecological and Cultural Barriers to Treatment of Childhood Diarrhoea in Riverine Areas of Ondo State, Nigeria. Social Science and Medicine 50 (2000) 953-964.

Geweke, J., Zhou, G. (1996). Measuring the pricing error of the arbitrage pricing theory. Review of Financial Studies 9, 557-587.

Gibson RC, Hotz C: Dietary diversification / study in Tanzania. Eur J Clin Nutr 2000;54: terlow JC (ed): Linear Growth Retardation in modification strategies to enhance micronutrient content and bioavailability of diets in developing countries. Br J Nutr 2001; 85: S159-S166.

Gibson, J. and Rozelle, S., 1998. Results of the household survey component of the 1996 poverty assessment for Papua New Guinea, Department of Planning, Port Moresby.

Gibson, J. 2000a. The economic and nutritional importance of household food production in PNG, in Bourke et al. 2000.

Gibson, J. 2000b. The nutritional status of Papua New Guinea's population, in Bourke et al. 2000.

Gibson, J. (2001), "Literacy and Intra-household externalities", World Development, vol. 29, pp. 155-66.

Glas, C.A., W. (2001). Differential item functioning depending on general covariates. In Boomsma,A. ,van Duijn, M.A.J. , Snijders, T.A.B. (Eds.) Essays on item response theory, 131-148. Springer, New York.

Gutner, T. (199). The Political Economy of Food Subsidy Perform in Egypt. International Food Policy Research Institute, FCND Discussion Paper No.77. Washington, DC.

Haddad, L., Ruel, T. and Garrett, L. (1999). Are Urban Poverty and Under nutrition Growing Some Newly Assembled Evidence?.International Food Policy Research Institute, FCND Discussion Paper No.63. Washington, DC.

Hastie, T. and Tibshirani, R. 1990. Generalized Additive Models. Chapman and Hall, London. Kammann, E. E. and Wand, M. P. 2003. Geoadditive Models. Journal of the Royal Statistical Society C, 52, 1-18.

Hauser, R.M., Goldberger, A.S. (1971). The treatment of unobservable variables in path analysis.InCostner, H.L.(Ed.) Sociological Methodology 1971, 81-117. Jossey-Bass, San Francisco.

Heywood, H.B. (1931).On finite sequences of real numbers. Proceedings of the Royal Society A 134, 486-501.

Hobert, J. P. and Casella, G. (1996). The effect of improper priors on Gibbs sampling in hierarchical linear mixed models. Journal of the American Statistical Association 91, 1461-1473.

Holmes, L.B., Harvey, E.A., Kleiner, B.C., Leppig, K.A., Cann, C.I., Munoz, A., Polk, B.F.(1987).Predictive value of minor anomalies:

II. Use incohort studies to identify teratogens. Teratology 36, 291-297.

Jeoreskog, K.G. and Goldberger, A.S. (1975). Estimation of a model with multiple indica- tors and multiple causes of a single latent variable. Journal of the American Statistical Association 70, 631-639.

Kandala, N.B.(2002). Spatial Modelling of Socio-economic and Demographic Determinants of Childhood Undernutrition and Mortality in Africa. Doctoral Thesis, University of Munich, Germany.

Kandala, N.B., Lang, S., Klasen, S. and Fahrmeir, L. 2001. Semiparametric Analysis of the Socio-Demographic Determinants of Undernutrition in Two African Countries. Research in Official Statistics, Vol. 4 No.1:81-100.217.

Kamman, E.E and Wand M.P. (2001). Geoadditive models, Department of Biostatistics, School of Publicc Health,Havard University.

Klasen, S. (1996). Nutrition, Health, and Mortality in Sub-Saharan Africa: Is there a Gender Bias? Journal of Development Studies 32:913-932.

Klasen, S. (1999). Malnourished and Surviving in South Asia, better Nourished and Dying young in Africa: What can explain this puzzle? SFB 386 Discussion Paper No.214. University of Munich.

Kirk wood BR. 1991. Diarrhea In: Feachem RG, Jamison E, Editors. Disease and Mortality in sub-Saharan Africa. Oxford: Oxford University Press, p.134-55.

Knut H. and Harald H. (2002). Cross-sectional Study of Morbidity, Morbidity-Associated Factors and Cost of Treatment in Ngaoundere, Cameroon, with Implications for Health Policy in Developing Countries and Development Assistance Policy, BMC Inter. Health and Human Rights.

Langsten R, Hill K (1994). Diarrhoeal disease, oral rehydration, and childhood mortality in rural Egypt. J Trop Pediatr. 1994 Oct; 40(5): 272-8.

Lang for d, I.H.; Leyland, A.H., Rabash, J., Goldstein, H. (1999). Multilevel Modeling of the Geographical Distributions of Diseases. Journal of Royal Statistical Society, Series A (Applied Statistics) 48:253-268.

Lang, S., Adebayo S.B., and Fahrmeir, L., Steiner, W.J (2002). Bayesian Geoadditive Seemingly Unrelated Regression. SFB 386, Discussion Paper 300.

Lang, S. and Brezger, A.(2004). Bayesian P-splines. Journal of Computational and Graph ical Statistics 13, 183-212.

Lang, S.; Brezger, A.; Kneib, T (2005). BayesX: Analysing Bayesian structured additive regression models. Journal of Statistical Software, 14 (2005), Heft 11.

Lavy V., Strauss, J., Thomas D. and de Vreyer, P. (1996), "Quality of Health Care, Survival and Health Outcomes in Ghana", Journal of Health Economics, vol. 15, pp. 333-57.

Lee, S.-Y. and Song, X.-Y.(2004). Bayesian model comparison of nonlinear structural equation models with missing continuous and ordinal categorical data. British Journal of Mathematical and Statistical Psychology 57, 131-150.

Lopes, H.F., West, M. (2004). Bayesian model assessment in factor analysis. Statistica Sinica 14, 41-67.

Magadi, M.A.; Madise, N.J.; and Rodrigues, R.N. 2000. Frequency and Timing of Antenatal Care in Kenya: Explaining the Variations between Women of Different Communities. Social Sciences and Medicine, vol.51: 551-561.

Magadi, M.A. 1997. Status of Women and Infant Child Health in Kenya with Particular Reference to the High Mortality Zone in Nyanza Province, UAPS Study Report No.26.

McCullagh, P. and Nelder, J. A. (1989). Generalized Linear Models. Chapman and Hall, London.

McCulloch, C. E. and Searle, S. R. (2001). Generalized, Linear, and Mixed Models. Wiley, New York.

Menon, M. Ruel, T. and Morris, S. (2000). Socioeconomic Differentials in Child Stunting are Consistently Larger in Urban than in Rural Area. International Food Policy Research Institute, FCND Discussion Paper No.97. Washington, DC.

Muth´en, B.O.(1984). A general structural equation model with dichotomous, ordered categorical and continuous latent variables indicators. Psychometrika 49, 115-132.

Muth´en, B.O.(1989). Latent variable modeling in heterogeneous populations. Psychome- trika 54, 557-585.

Muth´en,B.O.(2002). Beyond SEM: General latent variable modeling. Behaviormetrika 29, 81-117.

Miller, P. and Hirschhorn, N. (1995). The Effect of Nutritional Control of Diarrhoea Diseases Program on Mortality: The Case of Egypt. Social Science Medicine. Vol.40, No.10, pp.s1-s30.

Miller P, Loza S, Terreri N, Nagaty A, Bayoumi M, Hirschhorn N, Gipson R. United Nations Children's Fund, Cairo, Egypt. Diarrhoea and mortality in Menoufia, Egypt. J Diarrhoeal Dis Res 1994 Dec;12(4): 293.

Moradi, A.(1999). Determinanten der Unterernhrung in Kenya, Zambia,und In dien. Master's Thesis, University of Munich.

Moradi, A. (2002). HEIGHT AND HEALTH OF WOMEN IN SUB-SAHARAN AFRICA AND SOUTH-ASIA 1950-1980. Department of Economic History, University of Tuebingen, Germany.

Moradi, A. and S. Klasen. (2000). The Nutritional Status of Elites in India, Kenya, and Zambia: An Appropriate Guide for Developing International Reference Standards for Undernutrition SFB 386 Discussion Paper No. 217. University of Munich.

Moustaki, I., Jeoreskog, K.G. and Mavridis, D.(2004). Factor models for ordinal variables with covariate effects on the manifest and latent variables:a comparison of LISREL and IRT approaches. Structural Equation Modeling 11, 487-513.

Moustaki, I. and Steele, F. (2005). Latent variable models for mixed categorical and survival responses,with an application to fertility preferences and family planning in Bangladesh. Statistical Modeling 5, 327-342.

Nelder, J. A. and Wedderburn R. W. M. (1972). Generalized linear models. Journal of the Royal Statistical Society Series A 135, 370-384.

Nyovani, J.M., Z. Matthews, and B. Margetts (1999). Heterogeneity of Child Nutritional Status between Households: A Comparison of six Sub- Saharan African Countries. Population Studies 53: 331-343.

Obeng Gyimah, S (2003). Interaction Effects of Maternal Education and Household Facilities on Childho Diarrhea in sub-Saharan Africa: The Case of Ghana. Journal of Health and Population in Developing Countries. 13 November, 2003.

Pelletier, D. (1998). Malnutrition, Morbidity, and Child Mortality in Developing Countries, In: United Nations (eds.) Too Yuong too Die: Genes or Gender? New York: United Nations.

Quinn, K.M. (2004). Bayesian factor analysis form mixed ordinal and continuous responses. Political Analysis 12, 338-353.

Raach, A.W. (2005). A Bayesian semiparametric latent variable model for binary, ordinal and continuous response. Dissertation, Department of Statistics, University of Munich.

Rue, H. (2001). Fast sampling of Gaussian Markov random fields with applications. Journal of the Royal Statistical Society B 63, 325-338.

Rue, H. and Tjelmeland, H.(2002). Fitting Gaussian Markov random fieldsto Gaussian fields. Scandinavian Journal of Statistics 29, 31-49.

Rue, H.and Held, L. (2005). Gaussian Markov Random Fields: Theory and Applications.

Ryland, S. and H. Raggers. (1998). Child Morbidity and Treatment Patterns. DHS Comparative Studies, No.27, Claverton, Maryland: Macro International.

Ravi, P.; Eliya, R.; Blanco-Vidal, C.; Nandakumar, A. K. (1995). The Distribution of Health Care Resources in Egypt: Implications for Equity. An Analysis Using a National Health Accounts Framework. Associate Fellow, Health Policy Programme Institute of Policy Studies. 99, St. Michael's Road Colombo 3, Sri Lanka.

Sammel, M.D., Ryan, L.M. (1996). Latent variable models with fixed effects. Biometrics 52, 650-663.

Sammel, M.D., Ryan, L.M., Legler, J.M. (1997). Latent variable models for mixed discrete and continuous outcomes. Journal of the Royal Statistical Society B 59, 667-678.

Sanuntar. S (2005). Child Malnutrition as a Poverty Indicator: An Evaluation in the Context of Different Development Inventions in Indonesia. ADB Institute Discussion paper No.21.

Spearman, C. (1904). General intelligence objectively determined and measured. American Journal of Psychology 15, 201-293.

Smith, L.and L. Haddad. (1999). Explaining Child Malnutritionin Developing Countries. IFPRI. Research Report No.111, Washington DC: IFPRI.

Smith,L.and L.Haddad.(2001). The Importance of Women's Status for Child Nutrition in Developing Countries. Mimeographed, IFPRI. Wash- ington DC: IFPRI.

Spiegelhalter D., Best N., Carlin B., and Van der Line A. (2002). Bayesian Measures of Models Complexity. Journal of the Royal Stat. Soc. B(64):1-34.

Stephenson, C. (1999). Burden of Infection on Growth Failure. Journal of Nutrition, Supplement, 534S-538S.

Stoszek, K. Engle, R. Abdel-Hamid, M., Mikhail, N., Abdel-Aziz, F., Medhat, A., Fix, D., Emerson, U., Purcell, R. and Strickland, G. (2006). Hepatitis E antibody Seroconversion Without disease in Highly Economic Rural Egyptian Communities. Transactions of the Royal Society of Tropical Medicine and Hygiene (2006) 100, 89-94.

UNICEF (2000). The State of the World's Children. New York: UNICEF.

Toroitich-Ruto C (1998). Understanding Adolescent Sexual Behavior and their Reproductive Health Needs in Kenya. Journal of Obstetrics and Gynaecology of Eastern and Central Africa. March 14 (1).

Toroitich-Ruto C (1998). The Correlates of Low Nutritional Status among Young School- Going Children in Kenya. 36. (4). International Journal of Health Education, UK. 120- 126.

Toroitich-Ruto C (1999). The Relationship between Reproductive Health and Family Nutrition in Africa. Workshop on Women's reproductive Health and Household Food Security in Rural Africa organized by the ECA Food Security and Sustainable Division the (FSSDD), Addis Ababa, October, 1999.

Toroitich-Ruto C and S Wasao. Factors Affecting Children's Nutritional and Health Status in Kenya: A District-Level Analysis. Paper presented at the Population Association Of America annual meeting in New York Marriot Marquis Hotel on March 25-27, 1999.

Van Ginneken, J.K. and A.W. Teunissen. (1992). Morbidity and mortality from diarrhoeal diseases in children under age five in sub-Saharan Africa. In E. van de Walle, G. Psion and M. Sala-Diakanda, eds., Mortality and Society in sub-Saharan Africa. Oxford: Clarendon Press.

Walter D.R.O. (2001). Child Morbidity in Kenya: Does Women's Status Matter, Paper Presented at the Canadian Population Society 2001

Annual General Meeting, Laval University, Quebec City, May 27-29, 2001.

Woldemical G. (2001). Diarrhea Morbidity among Young Children in Erithrea: Environmental and Socioeconomic Determinants. Jour. Health Pop. Nutr., 19(2):83-90.

WHO (1983). Measuring Change in Nutritional Status.Geneva: WHO.

WHO (1999). Infant and Young Child Growth: The WHO Multicentre Growth Reference Study.Executive Board:Implementation of Resolutions and Decisions EB105/Inf.doc/1.Geneva:WHO.

World Health Organization. (1996). Childhood Diseases in Africa: Fact Sheets, p.1-6.

Yassin, K. (2000). Indices and Sociodemographic determinants of Childhood Mortality in Rural Upper Egypt. Social Science and Medicine 51 (2000) 185-197.

Yoannes AG, Streateld K. Bost L. 1992. Child Morbidity Patterns in Ethiopa, J. Biosoc. Sci., pp.:143-55.

Zwinderman, A. (1997).Response models with manifest predictors. Invander Linden, W.J., Hambleton,R.K.(Eds.) Handbook of modernitem response theory, 245-256. Springer, New York.