SMALL AREA ESTIMATION TECHNIQUES: FOCUS ON UNDER-FIVE MORTALITY DATA IN UGANDA

By

Asiimwe John Bosco
B.Stat (Mak), M.Stat (Mak)

A Thesis Submitted in Fulfillment of the Requirements for the Award of the Degree of Doctor of Philosophy in Statistics of Makerere University

May, 2014

Declaration

I, Asiimwe John Bosco hereby declare that the work contained in this thesis is my original work and has not previously in its entirety or in part been submitted at any University for a Degree award.

Signature: ..... Date: .....

Asiimwe John Bosco
PHD CANDIDATE

Approval

This is to certify that this thesis has been submitted as a fulfillment of the requirements for the award of a degree of Doctor of Philosophy of Makerere University, with my approval.

Signature: ..... Date: .....
Peter Jegrace Jehopio (PhD)
Supervisor

Signature: ..... Date: .....
Leonard K. Atuhaire (Associate Professor)
Supervisor

Dedication

I dedicate this work to my parents, Mr Christopher Rwabutiiti and Mrs Imelda Rwabutiiti for a big role played in building my solid academic foundation.

Acknowledgements

This work would not have been possible without the cooperation and advice from a number of people who contributed to its formation in one way or the other. Foremost, I acknowledge my advisors Dr. P. Jegrace Jehopio, Associate Professor Leonard K. Atuhaire, Dr. Fabian Nabugomu and Professor Xavier Mugisha for their tireless effort in ensuring that this thesis takes shape.

Secondly, I would like to extend my sincere appreciation to Makerere University and in particular the School of Graduate Studies for funding this research and paying my tuition. Special appreciations go to the former Director School of Graduate Studies, Professor Eli Katunguka-Rwakishaya, the former Deputy Directors, Associate Professor Christine Dranzoa and Associate Professor George Nasinyama. I would also like to thank other former or current staff members at the School of Graduate studies including; Mr. Katega, Mr. Kagumba, Mrs Baguma, Amina and Christine for their cooperation in enabling me to access funding and registration for the course. I also highly appreciate the support from Alan in the School of Graduate Studies and the zeal he attaches to his work. Alan reminds me of my great teachers/school colleague (Mr Justus Nyarwa, Hon. Henry Basaliza Araali and Mr Adolf Kabagambe) who had the zeal in their work and I believe they contributed greatly to my academic progression.

I am grateful to the Macro International and the Uganda Bureau of Statistics for providing the various data sets that I used in this study.

I am also humbled by the support from staff and friends at Makerere University, School of Statistics and Planning for their moral, academic and material support they accorded me during the period of my

research. Special attribute go to the following personalities**;** Mrs. Agnes Ssekiboobo, Associate Professor Leonard Atuhaire and Prof. James Ntozi, Dr. Tom Makumbi, Dr. Bruno Ocaya and Dr. Rebecca Nsubuga for their positive contribution during my proposal and progress report presentations. I would also like to thank the former staff of REACH/SURE Project, College of Public Health, Makerere University particularly Mr. Delius Asiimwe and Dr. Harriet Nabudere for enabling me to present part of my findings to policy makers drawn from the health sector and PhD students from Universities in the East African region, Canada and Cameroon in September, 2010 whose comments enriched my report. I acknowledge the academic contribution from Dr. Anthony K. Mbonye, Dr. Phyllis Freeman, Dr. Anthony Robbins and other anonymous reviewers/editors for helping to review the manuscript I submitted to the Journal of Public Health Policy (Asiimwe, Jehopio, Atuhaire, & Mbonye, 2010) as part of this research work. To my colleagues Mr.Abraham Owino, Mr. Christian Kakuba, Mr. Mike Barongo, Mr. Felix Wamono, Dr Robert Wamala, Dr Kizito Omalla, Mr. Charles Lwanga, Mr. Cyprian Misinde, Mr. John Mushomi, Ms Patricia Ndugga and Ms. Olive Buhule, I thank you for continuously encouraging each other. I further extend my gratitude to other members within the same School fo Statistics and Planning for the support they accorded me; Ms Esther Ngabirano (the School Registrar), Ms Hope Katushabe, Ms Rosette Nakawesa, Jimmy and Mariam. I am also grateful to the staff of Makerere University Library most specifically Ms Hellen Byamugisha for helping me to procure the most important text books for my research. I do also appreciate the staff from the College of Computing and Information Sciences who invited me to make presentations at their college and whose contributions enabled me to improve this research report. I feel extremely humbled by the great contribution from Professor Jon N. K. Rao the author of the book in the line of my research (Rao, 2003) together with Professor Jackie Galpin both initiated me into small area estimation technique, after attending Rao's

training for two weeks in Gauteng Province, Muldersdrift (Misty Hills Hotel), South Africa in November, 2007 at the Statistics South Africa (SASA) 50th Annual conference.

I also received a lot of personal support from the following South Africans when I visited their country in September 2007 while looking for opportunities to start my PhD and these include; Paul, Mr Trevor Manuel Bafana and some of his family members.

I could have probably mentioned my family first; Ronnie, Trisha, Ryan, Noah and Jesse for enabling me to pursue my studies and special thanks go to my wife, Mrs. Jennifer Kasabiiti Asiimwe. To the little peaceful babies Cristal and Aaronna, you were indirectly supportive to allow me have a good sleep after long hours of my PhD research work. I was always encouraged by my siblings, Immy, Liz, Rose, Dinah, Helen, Julius, Asumpta and Charles for continued positive inquiry into my academic progress.

List of Acronyms

AIDS	-	Acquired Immune Deficiency Syndrome
BYM	-	Besag, York and Mollie
CAR	-	Conditional Autoregressive
CREHP		Community Reproductive Health Project
CV	-	Coefficient of Variation
DIC	-	Deviance Information Criteria
DISH	-	Delivery of Improved Services for Health
EA	-	Enumeration Area
EB	-	Empirical Bayes
HB	-	Hierarchical Bayes
HMIS	-	Health Management Information System
IDP	-	Internally Displaced Persons
MC		Monte Carlo
MCMC		Markov Chain Monte Carlo
MoH	-	Ministry of Health
PSU	-	Primary Sampling Unit
RR	-	Relative Risk
SAE	-	Small Area Estimation
SD	-	Standard Deviation
SMR	-	Standardized Mortality Ratio
UBOS	-	Uganda Bureau of Statistics
UDHS	-	Uganda Demographic and Health Survey
URTI	-	Upper Respiratory Tract Infection
USAID-		United States Agency for International Development

Abstract

In Uganda, using survey data, estimates of under-five mortality have only been available at national and regional levels. This study utilized small area estimation techniques in a Hierarchical Bayes framework to derive estimates of relative risk of under-five mortality up to District level. The study utilized the Uganda Demographic and Health Survey data of 1995, 2001 and 2006 in the investigations. Results show that the Poisson-gamma model could provide reliable estimates for relative risk of under-five mortality. Results reveal that compared to the modeling approach, utilization of the traditional Standardized Mortality Ratio (SMR) could potentially be associated with very high undesirable coefficient of variations (>100%). The modeling approach has added advantage over the commonly used SMR by estimating under-five disease risk for a particular district and smoothening using adjacent district estimates. The study further reveals that it is possible to utilize small area estimation techniques together with national survey data to generate relative risk of under-five mortality for districts in Uganda. These results are potentially useful for targeting District decentralized system level of governance with high relative risk of under-5 mortality. **Key words: under-five mortality, district, small area estimation, disease mapping, Poisson-gamma, log-normal, Uganda.**

Table of Contents

List of Tables

List of Figures

Definition of Key Terminologies used in Study

Child mortality ($_4q_1$) is the probability of dying between exact ages one and five.

Disease Mapping is a geographical distribution of a disease within a population (Lawson & Williams, 2001).

Empirical Bayes methods are procedures for statistical inference in which the prior distribution is estimated from the data usually via the marginal distribution. This approach stands in contrast to standard Bayesian methods, for which the prior distribution is fixed before any data are observed.

Gibbs sampling is an algorithm to generate a sequence of samples from the joint probability distribution of two or more random variables. The purpose of such a sequence is to approximate the joint distribution, or to compute an integral (such as an ***expected value***). Gibbs sampling and the Metropolis-Hastings algorithm, are thus examples of a **Markov chain Monte Carlo** (MCMC) algorithm. Gibbs is one of the commonest methods of obtaining samples of parameters from posterior distributions, $f(\theta|y)$ that is estimation of relative risk (θ) given the data (y).

Bayesian Hierarchical Models or framework have an inherent hierarchical structure in which estimation of relative risk given the data $f(\theta|y)$ is estimated from the product of the likehood $f(y|\theta)$ and the prior $f(\theta)$. Let α and β be known parameter for instance in a gamma distribution. The prior on the other hand, is characterized levels of hierarchy; for instance $f(\theta|\alpha)$ at first level and $f(\alpha|\beta)$ at the second level. The structure can be terminated at simply the first level or extended to have more levels of hierarchy if required and hence a Bayesian Hierarchical model or framework.

Incidence is the number of new cases of disease which occur in a specified time period and in a specified population.

Infant mortality (1q0) is the probability of dying between birth and the first birthday.

National Surveys refers to surveys targeted to provide estimate or statistics for the entire country.

Neonatal mortality (NN) is the probability of dying within the first month of life.

Post neonatal mortality (PNN) is the probability of dying between the first month of life and first birthday (computed as the difference between infant and neonatal mortality).

Prevalence is the number of new and existing cases of disease which occur in a specified time period and in a specified population.

Prior probability distribution, often called simply the prior, of an uncertain quantity *p* is the probability distribution that would express one's uncertainty about *p* before the "data" e.g an opinion poll before a final vote.

Relative Risk is the ratio of the incidence of disease in the exposed population to the incidence in the non-exposed population (Lawson & Williams, 2001). It is a ratio of two probabilities.

Small area estimation is statistical techniques involving the estimation of parameters for small sub-populations, generally used when the sub-population of interest is included in a larger survey.

The posterior distribution is the product of the likelihood and the prior distribution and all the inference about parameters are made from the posterior distribution.

Under-five mortality ($_5q_0$) is the probability of dying between birth and the fifth birthday.

CHAPTER ONE: INTRODUCTION

This chapter provides background to the study; motivation for undertaking the study, problem statement, provides the scope, research objectives and discusses significance of the study. The chapter also states the limitations of the study and the ethical considerations. The study utilized both research hypothesis and research questions. For the latter, there were no test statistic used but graphical approaches were utilized.

1.1 Background to the Study

Small area estimations are statistical techniques involving the estimation of parameters for small sub-populations, generally used when the sub-population of interest are included in a larger survey. It is important to note that studies and literature on small area estimation that specifically targets under-five mortality are categorized as disease mapping. This study used small area estimation terminology instead of its subset 'disease mapping'. Small area estimation is a broader subject that provides estimates from a sample data for sub-population whose sample size provides unreliable estimates using direct estimation techniques. On the other hand, disease mapping is a sub-set of small area estimation that provides estimation mainly focusing on disease counts like the under-five mortality and morbidity.

Small area estimation involves using statistical models to link survey variable(s) of interest, such as disease, under-five mortality, poverty, etc to a local area such as a district or a sub-county whose estimates, because of small sample size at that level may not be derived by direct estimation from the available survey data. Small area estimates may be useful for government

agencies to allocate resources or identify hazardous areas related to high under-five mortality so that appropriate action may be taken (Lawson, et al., 2000; Meza, 2002; Wakefield, 2007). Mapping mortality and disease rates to display geographic variability is an increasingly common epidemiological tool and falls under a broad subject of small area estimation. Understanding spatial clustering of under-five mortality can provide a guide in targeting interventions in a more strategic approach to the population where mortality is highest and the interventions are most likely to make an impact (Lutambi, Alexander, Jensen, Mahutanga, & Nathan, 2010).

National surveys are widely used to provide estimates for the entire population parameters of interest but also for sub-populations (domains) such as regions, rural or urban, sex and age groups. However, such sub-populations are generally too large to provide a sense of particular lower level localities (small areas) like district or counties or sub-counties where the actual problem can easily be identified. Intervention can easily be accomplished when a small locality has been identified with a particular problem. Small area estimation provides a solution to using survey data to furnish estimates at such lower localities. The idea is that small area estimation techniques in particular "borrow strength" by using values of the variable of interest, y_i *(i=1,…D;* *D=Districts)*, from related areas to increase "effective" sample size. The value of, y_i, is by itself "too small" to provide a reliable direct estimate for a particular locality. For example, the number of under-five deaths, y_i, derived from a national survey data may be too small to provide an estimate of under-five mortality for a particular district. However, using small area estimation, the value of, y_i, can help derive a reliable estimate for the relative risk of under-five mortality for the district. In utilizing small area estimation techniques, the values, y_i, are brought into the estimation process through a model that provides a link to the related areas.

Since many health data are unavailable at the district level, policy makers sometimes rely on national-level datasets to understand the health needs of their communities (Jia, et al., 2004). Jia et al., also show that obtaining district level estimates can be accomplished using small area estimation techniques. Policy makers in Uganda equally use regional estimates to generalize values for the districts. National Survey data in Uganda like the demographic and health survey data provide direct estimates up to regional level. In this case the country was clustered into regions (Figure 3.1) and this implies that the sample can be sufficient to derive estimates (direct) up to regional level. The sample at the district level from the survey data would actually be "small" but using small area estimation techniques, reliable district level estimates like relative risk of under-five mortality can actually be derived. The term "small" therefore is derived from the fact that the sample for instance up to district level is generally small or will provide unreliable estimate up to that level. Small area is therefore a relative description of sample data that is not representative of a given locality or area.

1.2 Motivation for the Study

There were many cases when students and other data users were always tempted to generate direct estimates for the district level statistics from national surveys. This motivated me to pursue this study to demonstrate that for one to use district level data from the current national survey, there is need to do more in terms of modelling.

The area of spatial statistics appears to be gaining a lot of interest both locally and internationally. Locally, in the country many organizations I interacted with appear to be interested in displayed results on a map whenever services are offered in different areas like sub-

counties, districts or regions. Results displayed on maps appear to be highly attractive to readers and could easily attract the attention of policy makers who may not have adequate time to read reports. This scenario, to a great extent also motivated me into small area estimation techniques.

I have published at least three papers in international peer reviewed journals that are related to health particularly on mortality and this also created impetus to pursue a PhD in the same line of study.

1.3 Problem Statement

In Uganda, like in many developing countries often with budgetary constraints, censuses are usually carried out after every ten years to provide data up to the lowest administrative level like sub-county. However, censuses are usually limited in the variables captured due to the large coverage involved and the time lag of usually ten years. Other sources of data on mortality, like the Health Management Information System (HMIS) do not provide all cases of deaths since some of them that occur at household level are not reported to the health systems.

National surveys supplement censuses and are carried out at shorter intervals of five years. In Uganda, demographic and health surveys for instance are carried out regularly, usually after every five years. These surveys help in monitoring the trend in the country's performance on a number of indicators for instance those related to health and poverty. However, the sampling allows estimates only to be derived for large regional clusters and at times with few over-sampled areas in the country. Using national surveys, direct estimates like under-five mortality at district level are considered to be unattainable. For instance, the Demographic and Health Survey

2006 provides direct estimates for nine regions and for the over sampled areas of Karamoja[1] and the internally displaced camps in the northern region of the country. In the era of decentralization, especially at the district level, estimates for under-five mortality from such surveys using direct estimates are not achievable except by the application of small area estimation. Estimates like under-five mortality for the lower domains including the district level can only be achieved through use of small area estimation techniques which are usually avoided due to the advanced statistical modeling techniques that are involved. In this study, small area estimation techniques were used in a hierarchical Bayesian framework to show that one can derive relative risk of under-five mortality for districts in Uganda.

The conventional way to estimate relative risk of under-five mortality is to use the standardized mortality ratio (SMR) across different geographical districts. Direct use of SMR may not be worthwhile, particularly for small areas like districts (Bailey, 2001; Julious, Nicholl, & George, 2001b; MacNab, Farrell, Gustafson, & Wen, 2004; Meza, 2002), as it does not take into account the high variability for different population sizes over different regions and the spatial patterns of the regions under study. Use of SMR is based on ratio estimators and hence can yield large changes in the estimate, with relatively small changes in expected value (Downing, Forman, Gilthorpe, Edwards, & Manda, 2008; Lawson, Browne, & Rodeiro, 2003; Lawson & Williams, 2001). The use of hierarchical Bayes approach in small area estimation has been highly recommended as it smoothes the relative risks and provides the measures of uncertainty associated with these estimates of relative risk and the modeling can take into consideration the

[1]Area characterized by socio and economic hardships due to cattle rustling and unpredictable and hostile weather conditions

spatial autocorrelation. The approach to smoothing in the hierarchical Bayesian approach is by borrowing strength values from geographically referenced neighboring values.

In Uganda, estimates of under-five mortality have been only at national and regional levels. Reliable estimation of district parameters from survey data can accurately be achieved using small area estimation techniques. Due to the laborious and high statistical modeling that is involved in estimating relative risk of under-five for the district, these methodologies are usually avoided and policy makers tend to rely on regional estimates.

1.4 Scope of the Study

This study focused on small area estimation techniques that derive best predictor of relative risk of under-five mortality using a hierarchical Bayes framework. The study also compared the hierarchical Bayes framework with the SMR approach. The study utilized UDHS data sets of 1995, 2001 and 2006 to obtain variation over this period of time.

1.5 Research Objectives

The general objective of the study was to demonstrate the effectiveness of SAE techniques; particularly using under-five mortality data from Uganda. Specifically the study aimed:

i. To demonstrate the accuracy of a hierarchical Bayes framework, over the traditional SMR techniques.
ii. To establish a suitable model taking into consideration spatial autocorrelation that brings best predictor of relative risk of under-five mortality.

iii. To determine variation of relative risk of under-five mortality in Uganda, at the district level over the years 1995 to 2006.

iv. To determine areal pattern on relative risk of under-five mortality

1.6 Research Hypothesis

i. The SMR and the model approach using a hierarchical Bayes framework give the same results for the relative risk of under-five mortality

1.7 Research Questions

The research questions advanced in the study were the following

i. Does spatial autocorrelation have an effect in determining relative risk of under-five mortality?

ii. Are there no variations for relative risk of under-five mortality over a period of years 1995 to 2006 in Uganda?

iii. Are there spatial pattern for the relative risk of under-five mortality in Uganda?

1.8 Significance of the Study

This study demonstrates that modeling approach through use of the log-normal, Poisson-gamma and the BYM models yield better and more stable results than the traditional SMR. Although Maiti's study of 1998 is closely related, the study did not utilize the Poisson-gamma model and did not make a comparative analysis with the SMR. The Maiti study makes criticism about the use of SMR only based on literature but the results do not show to what extent HB results are better than those derived from the SMR. This study also demonstrates the use of small area estimation techniques to derive relative risk of under-five mortality for districts in Uganda. The

study produced outputs that are handy to districts to address under-five mortality. The study demonstrates the use of the national survey data by applying small area estimation technique to derive relative risk of under-five mortality up to district level. Districts with high relative risk of under-five mortality can be targeted for increased resource funding and mobilization. Furthermore, the study demonstrates that with the increase in the number of districts (from 37 in 1995 to 56 by 2006 in the survey data) and yet the sample size has not substantially been increasing, use of SMR becomes extremely unstable compared to the use of HB model approach which borrow strength from the neighboring districts to smoothen the estimate on relative risk of under-five mortality. The study provides a code fragment in WinBUGS 1.4 software that can be re-used in similar national surveys for Uganda or other countries.

The study helped to show the variation for the period 1995 to 2006 for the relative risk of under-five mortality at district level in Uganda though analysis was carried out independently for each year due to limited coverage as a result of insecurity for survey data of 1995 and 2001.

1.9 Limitations of the Study

The study used data from national surveys that have inherent gaps such as lack of data on children for women who had died though attempts were made to address them arising from the fact that only surviving women aged 15-49 years were interviewed. However, since the inherent problems are likely to be uniform across different districts, their effects on relative risk can be assumed to be uniform as well and less likely to affect relative risk of under-five mortality estimates.

Due to varying structure in the various datasets used in this study, a trend analysis in one model could not be achieved. For instance, some districts in 1995 and 2001 DHS were not covered due

to insecurity in some parts of the country. The data set for 2006, however, covered the whole country and hence was more reliable compared to the 1995 and 2001 DHS dataset. To this effect, analysis was done by looking at different datasets independently to show variation over time rather than obtaining trend.

1.10 Ethical Issues

This study was approved by Makerere University School of Graduate Studies. Permission to use the raw data was obtained from Macro International.

CHAPTER TWO: LITERATURE REVIEW

This chapter provides literature on studies done in relation to small area estimation techniques and those related to under-five mortality. The literature focuses mainly on studies done in the areas of under-five and child mortality and related to small area estimation. The chapter concludes with the conceptual framework used in the study.

2.1 SAE Studies Related to Under-five Mortality

Clayton and Kaldor (1987) applied Empirical Bayes (EB) estimation to data on observed case, y_i, and expected cases, E_i, of lip cancer registered during the period 1975-1980 in 56 counties (small areas) of Scotland. They reported the direct estimates, Standardized Mortality Ratio (SMR), the EB of θ_i based on the Poisson-gamma model and the approximate EB estimates of θ_i based on the log-normal model and the CAR-normal model for the 56 counties. Their findings indicate that SMR-values had a very wide variability compared to the other two EB estimates across counties. Leyland & Davies (2005) compared the Empirical Bayes methods with full Bayes methods for small area and their conclusion was that both methods have their place. Law, Serre, Christakos, and Leone (2004) carried out a spatial analysis and mapping of sexually transmitted diseases to optimize intervention and prevention strategies in North Carolina. Law, et al. (2004) used a simple Kriging geostatistical technique and argues that the methodology produces disease estimates with minimal mean square errors. Their study also points out that various types of kriging have been used in past epidemiological research work including simple, ordinary, and intrinsic kriging.

A study by MacNab et al., 2004 evaluated different approaches for relative risk inference to include the following; hybrid MCMC for Bayes analysis, type III parametric bootstrap, the carlin-gelfand sample reuse method and the penalize quasi-likehood (PQL) method for empirical Bayes. Their paper shows that the PQL is easy to implement and is computationally efficient.

Rao (2003) shows that small area estimation techniques can be used on mortality and disease rates for a given region or a county. Such maps are used to display geographical variability of a disease and identify high-rate areas warranting intervention. A simple small area model is obtained by assuming that the observed small area count, y_i, are independent Poisson variables.

Gangnon and Clayton (2003) proposed a Bayesian approach for inferencing the parameters of a hierarchical model of spatial clustering. In this approach, the disease rate of each region was explained through a combination of non-spatial fixed effects, spatial clustering effects, and spatially unstructured random effects. Here the non-spatial fixed effects component of the model consisted of a single parameter relevant to the overall rate across the study region, and the random effects were assumed to follow a normal distribution. The authors used the gamma prior distribution. Gangnon and Clayton paid most attention to modeling the spatial clustering effects, i.e. the relative risk for each region, which is the sum of log relative risks for potential clusters for each region.

One limitation of the Behavioral Risk Factor Surveillance System (BRFSS) data collected in the USA, is that valid estimates can only be obtained for states and larger geographic regions. Limited health data are available on the county level and, thus, many have used small-area

analysis techniques to estimate the prevalence of disease for the county level using BRFSS data (Goodman, 2010; Knutson, Zhang, and Tabnak, 2008). Jia, Muennig and Borawski (2004), used three methods of SAE including the synthetic method, spatial smoothing, and regression using the BRFSS data. The three small-area analysis methods were then applied to 2000 BRFSS data to examine how well each technique predicted county-level disability prevalence. Their results show that regression method produced the most valid and precise estimates of county-level disability prevalence though Kleinschmidta, Bagayokob, Clarkec, Craiga, and Le-Sueura, 2000 tend to disagree with the method when used alone, that results have a tendency to produce predicted values that are pulled towards the mean. Schneider, Lapane, Clark, and Rakowski (2009) equally used the same dataset with synthetic method to derive small area estimates for mammography, a study of the breast using x-ray, for women aged 40 to 79 years up to county level. Schneider et al. explains that the method allows documentation of geographic disparities and improves understanding of the spatial distribution of mammography prevalence. In yet another closely related study, Knutson, et al., (2008) used small-area estimation method by including individual and community data in a generalized, linear, mixed-effect model in a breast cancer study. In a cancer study, using SAE techniques, Maiti (1998) used the generalized linear mixed model (Logistic regression). In Maiti's paper, he also considered a CAR spatial model to generate estimates for lip cancer incidence in Scotland for each of 56 counties.

In a study by Alexander, Moyeed, and Stander (2000), the authors looked at individual-level counts of nematodes, a parasite of humans which causes the disease lymphatic filariasis. They used the negative binomial distribution with the argument that the shape parameter is a convenient index of over-dispersion. The negative binomial model is supported by Anderson

(1993) for which parasite counts are usually over-dispersed relative to the Poisson distribution. A related study by Lord (2006), show that a low sample size can potentially affect results from a Poisson-gamma model.

Adebayo and Fahrmeir (2005); Khatab and Fahrmeir (2009), used a Bayesian approach to analyse child mortality in Nigeria and Egypt respectively using geoadditive models. The application was based on the Demographic and Health Survey data in the respective countries. Other related studies include; Adebayo, Fahrmeir, and Klasen (2004); Kandala, Ji, Stallard, Stranges, and Cappuccio (2007). In these studies, the strength lies in the fact that they took into account the time element of the age at deaths of the child and included covariates in their models. Adebayo and Kandala studies however, provide results more for the regional clusters than actually the district (state) specifics. Adebayo's attempt to show district level burden of child mortality is merely based on relative frequency, a ratio of observed deaths to the total number of child which is more of SMR.

Empirical Bayes methods in particular have been criticized in favour of the Hierarchical Bayes method for not providing measure of variance of the relative risk estimate (MacNab, et al., 2004). However, Maheswaran, et al. (2006); Zhu, Carlin, English, and Scalf (2000) supports both the Hierarchical Bayes and empirical Bayes methods with a view that they have proven effective in smoothing crude maps of disease risk, eliminating the instability of estimates in low-population areas while maintaining overall geographic trends and patterns. Zhu, et al. (2000) extended the Hierarchical Bayes method which relates traffic density to pediatric asthma hospitalizations (disease counts- Y_i) with basic idea of modeling the number of disease events in region i, Y_i , as a Poisson random variable. Similarly, the Poisson Bayesian hierarchical model

was used by Haininga, et al. (2010) in estimating the relative risk of stroke from air pollution exposure using small area data.

A study by Ramis-Prieto, et al. (2007) compared models based on the Bayesian hierarchical models for modeling of Municipal mortality due to hematological neoplasias in Spain. The models included the Besag York and Mollie (BYM) model and a model based on zero-inflated Poisson (ZIP) distribution. They used the Deviance Information Criteria (DIC) to test for the goodness of fit. Their results showed that the models actually yielded similar results. The selection of the better model with a lower DIC was also used by Earnest, et al. (2010) in a similar study on small area estimation of sparse disease counts. A comparative study between the Poisson kriging and the BYM model for mapping disease risks show that more attention should be paid to the spatial and distributional assumptions underlying the BYM model. The Poisson kriging offers more flexibility in modeling the spatial structure of the risk and generates less smoothing, reducing the likelihood of missing areas with a high relative risk (Goovaerts and Gebreab, 2008).

Lawson, et al. (2000) evaluated the goodness-of-fit for the various small area methods and their finding show that the gamma-Poisson exchangeable model and the BYM were the most robust across a range of diverse models. The mixture models were less robust while the non-parametric smoothing methods performed badly. It was also concluded in their study that the linear Bayes methods displayed similar behavior as that of the Poisson-gamma methods.

Waldhoer, Wald, and Heinzl (2008) used the multivariate modeling of indirectly SMRs using the MCMC methods implemented in WinBUGS software based on observed and expected counts in

the analysis of spatial distribution of infant mortality by cause of death in Austria for the period 1984 to 2006. Important to note in their study is the fact that two adjacent districts which had small number of deaths were combined in the analysis and such aggregation is also supported by Arul, et al. (2007). Waldhoer, et al. (2008) used the conditional autoregressive model (CAR) also known as a BYM or a convolution model introduced by Besag et al. They argued that this model, assumes that the observed number of counts in spatial unit *i* is Poisson distributed with expectation μ_i. One reason for its popularity is that it is a straightforward estimation by the MCMC techniques e.g. using WinBUGS software. The BYM model is supported by Arul, et al. (2007); Downing, et al. (2008); Earnest, et al. (2007); Goovaerts and Gebreab (2008) for it smoothes relative risk estimates in each region towards the mean risk using the neighboring areas. They argue that the BYM model provides for a more precise or reliable estimate of both mean and variance compared to using the crude rate of SMR. This is especially so, as the variance for the estimate of the raw rate with a small expected count can be large and unreliable under the crude SMR. Risks are also smoothed towards the global mean to account for overdispersion caused by unobserved confounding factors (Bailey, 2001; D. Clayton and Bernardinelli, 1996; Mollie, 1995). Similar studies in using the CAR and Poisson spatial models were utilized by López-Abente, et al. (2008) on individual death entries for the period 1989–1998 corresponding to kidney cancer mortality in Spain, broken down by town or city, nationwide. SAS (2010) attributes overdispersion to positive correlation among the observations, an incorrect model, an incorrect distributional specification, or incorrect variance functions and shows that the Bayesian hierarchical Poisson regression models are effective in capturing overdispersion and provides a better fit.

A study by Downing, et al. (2008) modeled jointly the incidence rates of six smoking related cancers in the Yorkshire region of England, to explore the patterns of spatial correlation amongst them, and estimated the relative weight of smoking and other shared risk factors for the relevant disease sites, both before and after adjustment for socioeconomic background. Their conclusion, indicate that incidence estimates are more precise than those obtained without smoothing. A similar study by Randramanana, Richard, Rakotomanana, Sabatier, and Bicont (2010) used a Bayesian approach, to measure the associations between the spatial variation of TB risk and the national control program indicators for all neighborhoods in Madagascar. Randramanana et al., recommends the use of spatial approaches for assessing the epidemiological situation for TB.

2.2 Use of Standardized Mortality Ratio

Use of SMR has however, been subjected to a number of criticism with the fact that they can yield very large values for any positive count (Lawson, et al. 2003). Approach to the improvement of the relative risks estimates when SMR are used, is to employ smoothing tools to reduce the noise. One such approach is to use small area estimation techniques in a hierarchical Bayes framework by pooling information from the geographical referenced neighboring areas (Graham, 2008; A. B. Lawson, et al., 2003; Ocaña-Riola, 2007; Rao, 2003).

A study by *Julious, Nicholl, and George (2001a)* recommends that use of SMR should be avoided where necessary. A Comparative study that used a regression-adjusted mortality to standardized mortality ratios for trauma center (*Moore, Hanley, Turgeon, and Lavoie 2012)* reveals that the latter is an inferior method to the former and provides biased results. Other study (Symons and Taulbee, 1980) has shown that when the true relative risk is greater than 100%, the SMR over-estimates relative risk no matter how small the mortality rates are.

There are however, studies where SMR appears to have worked well (Kahn, Kramer, and Rubenfeld, 2007; Mok, Kwok, Ho, Chan, and Yip, 2011; Walsh, et al., 2012) though it is mainly recommended to derive relative risk for large geographical regions such as countries or states, but may be unreliable for small areas such as counties (Meza, 2003).

2.3 Methodologies for Estimating Under-five/Child Mortality

This section provides discussions on the methodologies available for estimation of under-five mortality as a key aspect in this study. The study concludes with reasons as to why traditional methodologies of estimating under-five mortality do no not provide estimates of up to district level in Uganda.

Under-five estimates in Sub-Sahara Africa have tended to use DHS. A study by Korenromp, Arnold, Williams, Nahlen, and Snow (2004) show that the DHS estimate in Sub-Sahara Africa on under-five mortality rates use complete birth histories, based on reports from mothers on the survival of their children. The direct estimation technique is based on a life table approach: probabilities of dying are computed from reported dates of birth and death and the numbers of children of a particular age exposed to the risk of dying during a specified period.

In countries with complete vital registration systems that capture all births and deaths, under-five mortality can be calculated using direct or indirect demographic techniques (Rajaratnam, N., Lopez, and Murray, 2010). In the absence of a complete vital registration system especially in developing countries, under-five mortality is estimated using censuses and surveys (Chowdhury, Islam, and Hossain, 2009). In Uganda, under-five mortality rates have been typically computed

using two approaches — direct and indirect techniques by utilizing censuses and surveys. For this study, direct estimates could be computed from the three Uganda Demographic and Health Surveys of 1995, 2001 and 2006 from data collected using the birth of a woman aged 15 – 49 years. However, due to lack of necessary data in all the three surveys, indirect techniques are a feasible alternative. The underlying assumptions used in the indirect methods can introduce a potential bias in the estimate. Studies have shown that for many Sub-Saharan countries, even if an appropriate mortality model is applied in the indirect estimation method, the results of this method are consistently higher than those of the direct methods (UBOS and Macro International, 2007). The best-known and most widely applied indirect estimation method was developed by William Brass in 1968. The method measures under-five mortality from dead children as a proportion of those ever born by women classified by age group (Hill and Figueroa, 2001). The basic principle of the method is that the age of the mother can serve as a proxy for the exposure time of her children, so that the proportion dead for women of a given age group can be converted into a defined probability of dying for their children. The method by William Brass has been applied to census and survey data from all parts of the developing world, and has been found to work remarkably well in a wide variety of settings (Hill and Figueroa, 2001).

The original method has been extended by a number of authors, notably Sullivan (1972), who included groups of women classified by age. In a modification of Brass's technique, Sullivan (1972) developed a simple linear regression model. Trussel (1975) later expanded the model base of the estimation methods (UN, 1983).

Although the Brass method is popular for estimation of under-five mortality, it has been subjected to criticism. One of the key assumptions underlying the Brass method is that under-

five mortality risks are uniform across the classificatory variable (age or duration of marriage of the mother) being used as a proxy for exposure to risk of the children. However, younger women aged 15-24 are almost always having higher estimates compared to older women. This pattern results from a real age effect, whereby children of young mothers have elevated mortality risks and also from a selection effect whereby women of lower socio-economic class tend to start childbearing early and have children exposed to above average mortality risk (Hill and Figueroa, 2001),

Using indirect techniques, the Trussel method which was developed in 1975 can be used in computing the under-five mortality. The Trussel method uses data classified by five-year age group of the mother. In particular, the number of children ever born, number of children surviving (or the number dead) and the total number of women (irrespective of marital status) can be considered in the computation of the under-five mortality. It is important to note that both the Brass, Sullivan and Trussel methodologies for determining under-five mortality using survey data cannot derive estimates up to district level where the sample size is not representative of those districts. This limitation therefore makes the study on small area estimation relevant. It is also important to note that under-five mortality is derived from the DHS data using direct estimation techniques and this study adopted these estimates.

The number of deaths (y_i), are usually based on a specified period of time – usually 5 or 10 years before the survey (WB, 2008). In this study to gain enough observations for the model a longer period was used for up to 35 years.

2.4 Conceptual Framework

The conceptual framework below is an illustration of the hierarchical approach used to estimate relative risk of under-five mortality (θ_i). To derive θ_i, the observed values y_i are necessary in other words we are estimating θ_i given the data (y_i) or simply expressed as $f(\theta_i|y_i)$, usually referred to as the posterior distribution. Let α and β be the parameters associated with the posterior distribution. The posterior can then be expressed in terms of the likelihood ($f(y_i|\theta_i)$) and prior distributions $g_1(\theta_i|\alpha,\beta)\ g_2(\alpha|\beta)\ g_3(\beta)$ as shown in equation 2.1.

$$f(\theta_i|y_i) \propto f(y_i|\theta_i)\ g_1(\theta_i|\alpha,\beta)\ g_2(\alpha|\beta)\ g_3(\beta) \quad \text{.....2.1}$$

At the first level of the hierarchy, $g_1(\theta_i|\alpha,\beta)$ and g_1 can be assumed to follow a gamma distribution. We can also have the second level ($g_2(\alpha|\beta)$) and the third level $g_3(\beta)$ but can be terminated at the first level (see Figure 2.1).

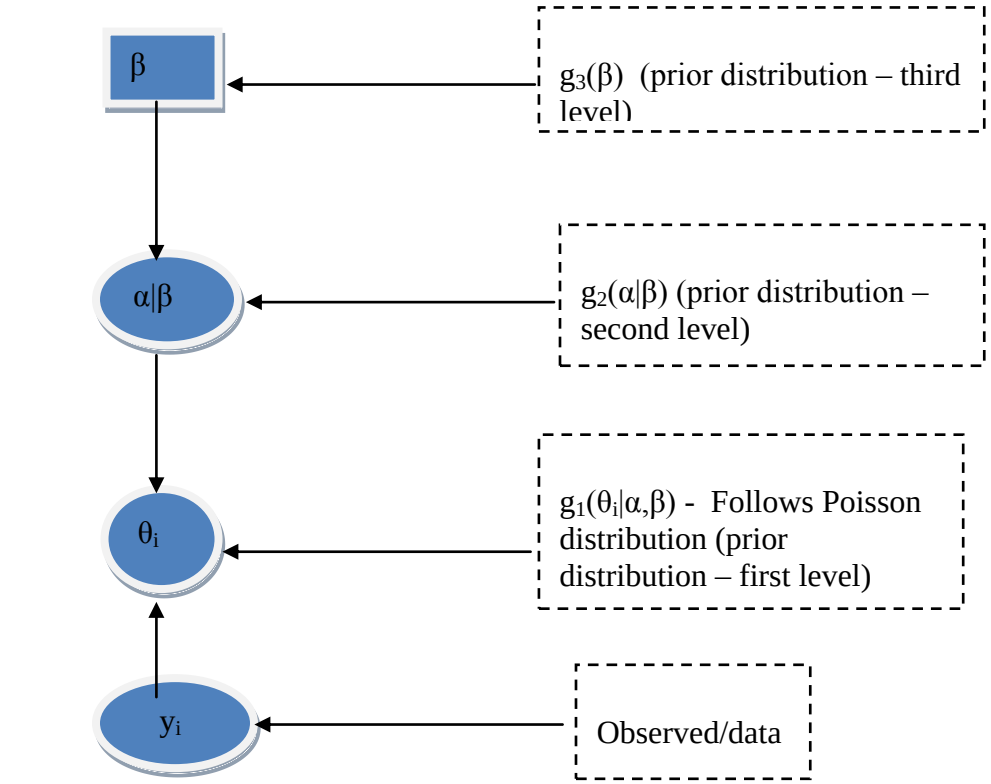


Figure 2.1: Two level hierarchical framework for θ_i

CHAPTER THREE: METHODOLOGY

This chapter looks at the methodologies that were used in the study, commencing with a brief description on the meaning of small area estimation. The chapter provides discussion on the data sets that were utilized. The chapter provides discussion on the commonly used method for estimating relative risk using the SMR. Various models used under the Bayesian Hierarchical framework (Poisson-gamma, log-normal and the Besang, York and Mollie Models) for estimation of relative risk of under-five mortality are also presented. The Bayesian Hierarchical models or framework is one of the various statistical techniques under the broader subject of small area estimation. The chapter provides discussion on the prior distributions that were used under the Bayesian Hierarchical framework. The chapter concludes with the ethical consideration and assumption made in the course of the study.

3.1 Data Description: Sources and Weakness

The study utilized data obtained from the Uganda Demographic and Health Surveys (UDHS) of 1995, 2001 and 2006. This section provides discussion on the three data sets that were used in the study; their sources and weakness.

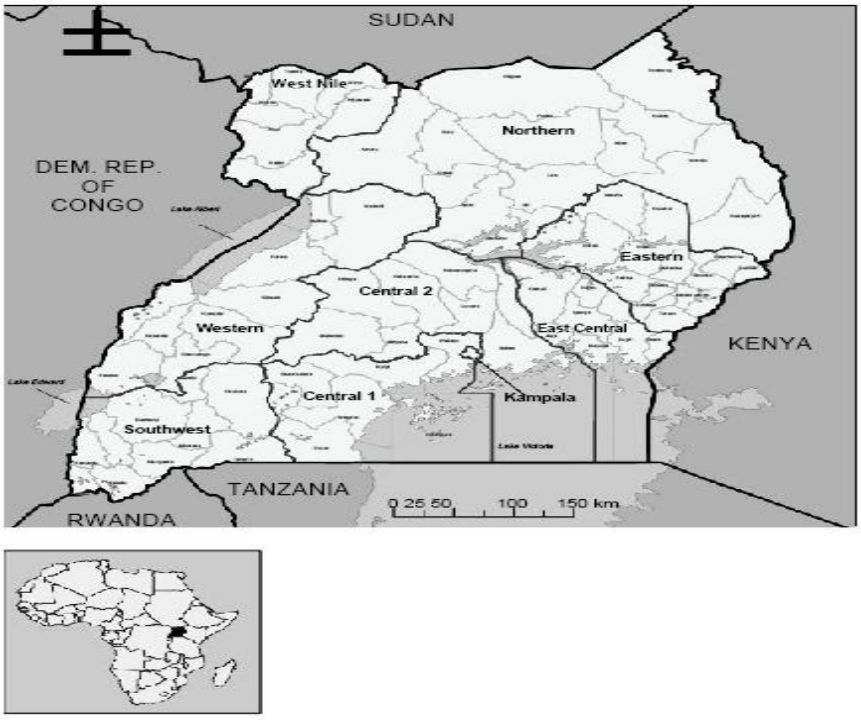
The UDHS survey of 1995 covered a total of 37 districts and due to insecurity, the district of Kitgum located in the northern part of the country was not covered. By the time of the survey, Uganda had a total of 38 districts. A sample of 303 Primary Sampling Units (PSUs) consisting of Enumeration Areas (EAs) were selected from a sampling frame of the 1991 Population and Housing Census and covered a total of 7,070 women in the reproductive age group of 15 – 49 years. The survey also obtained data from a total of 7,550 households and 1,996 men in a

reproductive age group of 15-54 years. The country was clustered into four regions consisting of Central, Eastern, Northern and Western. To permit calculation of contraceptive prevalence rates under a USAID-funded project called Delivery of Improved Services for Health (DISH) a sample design allowed for over sampling of households in the nine districts. These districts were Kasese, Mbarara, Masaka, Rakai, Luwero, Masindi, Jinja, Kamuli and Kampala.

The UDHS data of 2001 covered a total of 34 districts and again due to insecurity the districts of Kasese and Bundibugyo in Western Region as well as Gulu and Kitgum in the Northern region were excluded from the survey. A sample of 298 PSUs consisting of EAs were selected from a sampling frame of the 1991 Population and Housing Census and covered a total of 7,246 women in the reproductive age group of 15 – 49 years. The survey equally obtained data from a total of 7,885 households and 1,962 men in a reproductive age group of 15-54 years. The country was also clustered into four regions consisting of Central, Eastern, Northern and Western. To permit calculation of contraceptive prevalence rates under DISH project the nine districts were again over sampled. Over sampling of EAs was also carried out for the districts of Kabale, Kisoro and Rukungiri under the project called Community Reproductive Health Project (CREHP).

As compared to the UDHS of 1995 and 2001, the survey of 2006 covered all the 56 districts of the country providing a better estimate of relative risk of under-five mortality. It is important to note that the number of districts in the country have continued to increase to currently over 110. A sample size of 321 PSUs consisting of EAs were selected from a sampling frame of the clusters sampled in the 2005-2006 Uganda National Household Survey and an additional 47 EAs were over sampled from the North Eastern Region (Kotido, Moroto and Nakapiripirit) and IDP

camps in the districts of Gulu, Kitgum, Lira and Pader. The over sampled areas were mainly aimed at obtaining specific baseline indicators due to insecurity that ravaged the region for over 20 years prior to the survey. The country was clustered into nine regions compared to the four covered in the prior surveys. The nine regions included; Central 1, Central 2, Kampala, Eastern, East Central, North, West Nile, Western and South Western (See Figure 3.1). In general, a total of 8,531 women in the reproductive age group of 15 – 49 years were interviewed. The survey equally obtained data from a total of 8,870 households and 2,503 men in the reproductive age group of 15-54 years.



Source: UBOS, 2007

Figure 3.1: Map of Uganda showing clusters used in UDHS 2006

In each of the UDHS of 1995, 2001 and 2006 the main variable used in this study (under-five deaths) was collected from women in a reproductive age group of 15–49 years.

Data was accessed with permission from *MeasureDHS* website. This research work used under-five mortality rather than other mortality measures like infant or child mortality due to the fact that significantly larger and more samples are obtainable. Additionally, the indicator is in line with the MDGs target to reduce under-five mortality by two-thirds between 1990 and 2015. The data used in the estimation of the under-five mortality rates were collected on the birth history of women aged 15-49 years. For children who had died, the mothers were asked to provide the age at death. The data used for computation of under-five mortality is susceptible to some errors. Firstly, only surviving women aged 15-49 years were interviewed; therefore, no data are available for children of mothers who had died. Another possible error in data collection is underreporting of events (births and deaths), especially in cases where deaths occur early in infancy. Attempts to address under reporting of age at deaths were done by recording days if the death took place within one month after birth, in months if the child died within 24 months, and in years if the child was two years or older (UBOS and Macro International, 2007).

3.2 Use of SMR

3.2.0 Introduction

The SMR is defined as the number of observed deaths in the study population divided by the number of expected deaths (Last, 2001). If the SMR is greater than one (1) means the number of observed deaths is greater than what would be expected if the study population had the same probability of dying as the standard population, while an SMR of less than one (1) means the number of observed deaths is less than expected.

3.2.1 Computation of Relative Risk

There are two principal categories of estimation methods for calculating under five mortality rates: direct and indirect. The direct methods employ data on the date of birth of children, their survival status, and the dates of death or ages at death of deceased children. The indirect methods use information on survival status of children to specific age cohorts of mothers. The indirect methods can utilize data that are commonly collected in censuses and surveys: the number of children ever born and their age. However, direct and indirect estimation techniques of under-five mortality cannot provide estimates up to district level from available survey data. The simpler way is to use a SMR approach which has traditionally been used as an estimate of relative risk of under-five mortality. Let y_i and E_i denote the number of deaths and the expected number of deaths respectively from the disease during the study period. The Standardized Mortality Ratio (SMR) is defined as;

$$\hat{\theta}_i = \frac{y_i}{E_i} \quad \ldots\ldots 3.1$$

Generally the expected number of deaths E_i is assumed known (Bailey, 2001). Let r_j be the under-five mortality rate in the j^{th} region and p_{ij} is the population in the i^{th} district located in the j^{th} region. The formula for the computation of expected deaths is then given as;

$$E_i = \sum_j r_j p_{ij} \quad \ldots\ldots 3.2$$

3.2.2 Computation of Expected Value

Computation of expected value for the under-five was based on estimated population in age group below five years and the under-five mortality rate for a given region. To estimate the population below five years, this study utilized the census figures of 2002 together with the

exponential growth function. For instance, Adjumani district located in the North Western part of the country had a population of 202,290 by the year 2002. In 2006, Adjumani district had under-five mortality rate of 177 deaths per 1,000 live births. To estimate the population for the district by the year 2006, the following exponential growth rate function was used;

$$N_t = N_0 e^{rt} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.3$$

Where N_t is the population estimate for Adjumani for the year 2006 based on census figures of 2002 after a time period (t) of four years. A population growth rate (r) of 3.4 percent with about 19.2 percent of the population below age of five years (UBOS, and Macro International, 2007) was used. N_0 is the population census figure of 2002. Using equation (3.2) the expected deaths below the age of five in the district of Adjumani provides a figure of about 7,876 cases by year 2006 (See Appendix 5).

3.2.3 Computation of Standard Deviation, Standard Error and Coefficient of Variation for the Relative Risk

Let y_i and E_i refer to the number of observed deaths and expected number of deaths among under-five in a given district respectively.

The standard deviation (s.d) for the SMR is given as (Soe and Sullivan, 2006);

$$s.d_i = \frac{1}{\sqrt{y_i}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.4$$

The standard error (s.e) for the SMR is given as;

$$s.e_i = \frac{E_i}{\sqrt{y_i}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.5$$

Coefficient of variation is then defined as;

$$cv_i = \frac{s.d_i}{SMR_i} * 100 \quad \text{.....3.6}$$

3.3 Small Area Estimation Methodologies

This section looks at the different models used under the Hierarchical Bayes framework that were used in this study. The models discussed include; Poisson-gamma, log-normal and the Besang, York and Mollie (BYM). The section also provides discussion on the different priors that were used in the models.

3.3.1 Poisson-gamma Model

In a Bayesian setting, we have a random sample $y_1,...,y_n$ from a Poisson($E_i\theta_i$) distribution for counts events. Where $y_1,...,y_n$ are the under-five deaths from n districts while θ_i are the relative risks for $i=1,....n$ districts to be estimated. We are interested in obtaining the posterior. The posterior is the product of the likelihood and the prior distributions and all the inference about parameters are made from the posterior distribution i.e

Posterior $\propto$ prior . likelihood

Or

$$g(\theta_i|y_1,...,y_n) \propto g(\theta)\, f(y_1,...,y_n|\theta_i). \quad \text{.....3.7}$$

The likelihood of a Poisson distribution is a product of the original likelihoods which simplifies to:

$$\prod_{i=1}^{n} (y_i \mid \theta_i) \propto \theta_i \sum_{i=1}^{n} y_i e^{-n\theta_i} \quad \text{.....3.8}$$

A Hierarchical Bayesian Method was used due to the fact that in each area-specific prevalence an advantage of pooling information from neighboring areas or a prior distribution is utilized (Bergamaschi, et al., 2006). The Hierarchical Bayesian methods have an important role in modeling complexity of data structures in spatial epidemiology. The Bayesian approach in small area modeling consists of considering, in addition to the likelihood of the observed counts, prior information on the variability of the relative risks which subsequently is considered to improve the estimates from the posterior.

A Poisson model is considered a basic one with parameter $E_i\theta_i$ and is referred to as a classical model (SMRs). The other models like log-normal and the Besag, York and Mollie (BYM) model are extensions of this classical model (Lawson, et al. (2003). The Poisson-gamma model is suitable for non-contagious and rare cases, where the numbers of deaths in each area are assumed to be mutually independent and hence follow a Poisson distribution. It was therefore imperative for this study to begin with the basic classical model and then proceed to evaluate whether there is benefit in using other models.

The Poisson models in general are used when the dependent variable is a count of rare events like accidents and deaths among the population.

Let Y be the vector for the under-five deaths in districts. At the level of direct estimation, Y follows a Poisson distribution given by;

$$\Pr(Y = y) = f(y;\theta) = \frac{\exp(-\theta)\theta^{y}}{y!}.$$

……..3.9

For y=0,1,2,…, θ >0

Let *n* be the number of districts under consideration. Let θ_i be the relative risk of a disease for area *i (i=1,…,n)* in reference to the under-five mortality rate p_j in stratum *j* (region). The problem is to estimate θ_i *(i=1,…,n)* and its variance or standard deviation. Let y_i denote the observed number of death in district *i (i=1,…,n)* observed over a period of time. Let E_i be the expected deaths in area *i (i=1,…,n)* and considered to be known. y_i*'s* represent the observed values of under-five deaths and is assumed to be independent Poisson random variables with parameter ($\theta_i E_i$) and with mean μ and variance δ^2. Using the hierarchical models and the gamma distribution considered for θ_i then we have;

$y_i \sim \text{Poisson}(\theta_i E_i)$; independently, *i (i=1,…,n);* ……..*3.10*

given one set of data (dependent), *y (*referring to y_i*)* and similarly for the relative risk *(θ)* , then the general distribution in the hierarchical Bayes framework is given as;

$$f(\theta|y)=\frac{f(y\mid\theta)f(\theta)}{f(y)} \propto f(y\mid\theta)f(\theta) \quad \text{……..3.11}$$

That is the posterior is proportional to the product of the likelihood and the prior distribution. The likelihood is given as;

$$f(y_i\mid\theta_i)=\prod_{i=1}^{n}\frac{e^{-(\theta_i)}(\theta_i)^{y_i}}{y_i}=\frac{e^{-n(\theta_i)}(\theta_i)^{\sum_{i=1}^{n}y_i}}{\prod_{i=1}^{n}y_i!} \quad \text{……..3.12}$$

The prior distribution $f(\theta_i)$ is assumed to follow a Gamma distribution with parameter α and β>0 i.e $\theta_i \sim \Gamma(\alpha,\beta)$ then the probability density function is given as;

$$f(\theta_i\mid\alpha,\beta)=\frac{\beta^{\alpha}e^{-\beta\theta_i}\theta_i^{\alpha-1}}{\Gamma(\alpha)} \quad \text{……..3.13}$$

Where $\Gamma(\alpha) = \int_0^{\infty} \theta^{\alpha-1} e^{-\theta} \partial\theta$

Using equation 3.11, then the resulting posterior distribution is given by;

$$f(\theta_i \mid y_i) \propto f(y_i \mid \theta_i) f(\theta_i) \propto \frac{e^{-n\theta_i} \theta_i^{\sum_{i=1}^{n} y_i}}{\Pi_{i=1}^{n} y_i!} \frac{\beta^{\alpha} e^{-\beta\theta_i} \theta_i^{\alpha-1}}{\Gamma(\alpha)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.14$$

with the mean given by (Lawson, 2003);

$$E[\theta_i|y_i,\alpha,\beta] = \frac{\alpha + y_i}{\beta + E_i} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.15$$

One disadvantage of the Poisson-gamma model is its inability to cope with spatial correlation. Presence of spatial correlation in the model violets the assumption of independency with the Poisson-gamma model. Models that suitably takes into account spatial correlation are; log-normal and the BYM (Lawson, et al., 2000; Lawson, et al., 2003).

3.3.2 Log-normal Model

The basic model of the Poisson when assumed to be independently distributed is subject to a number of problems. These problems include the fact that spatial correlation within the districts can affect the sampling distribution under the standard Poisson model (Rodeiro and Lawson, 2004). Lawson (1994) proposed two approaches; first, a standard Poisson model could be assumed and the residual from the fit could be examined for the spatial autocorrelation. If the residuals are effectively independent and identically distributed (*i.i.d*) then independent assumption is tenable. The second approach is where there is pure autocorrelation in the data in which a model that includes autocorrelation could be admitted.

If spatial autocorrelation in the residuals exist, it means that the model is systematically overestimating the observed values in some districts, and underestimating the observed values in other districts. In addition, we will get unrealistic values for the significance and confidence limits for the coefficients, especially if we assume that the mortality in one district is independent of the other neighboring district. If there is correlation between them, they are not independent implying that model estimations will be biased.

The other assumption made in the basic Poisson model is that the mean and variance are equal. Sparseness of data or a large incidence of zero counts like in mortality data may bring in extra variation and may cause violation of the mean-variance assumption. In this case, extra Poisson model has been proposed where $\text{var}(\lambda_i) = \alpha E(\lambda_i)$ by introducing a factor α . Negative binomial or zero-inflated Poisson models have also been proposed in solving violation of Poisson model assumption (Lambert, 1992). In the hierarchical Bayes method, random effects modeling is often used to deal with the problem of over dispersion in modeling count data. A study by Catelan, Lagazio, and Biggeri (2010) show that the Poisson-gamma and BYM models can account for extra Poisson variability in a Bayesian formulation. Geographical locations like district treated as small area in this study, may exhibit some form of spatial autocorrelation. The fact that the districts are simply clusters that were created for administrative purpose, neighboring ones may share similar socio-economic characteristics like terrain nature that may have for instance similar water source problems and hence may eventually exhibit similar disease patterns that could be correlated with under-five deaths (y_i). In this study the effect of spatial autocorrelation that may exist within the geographical location like districts was evaluated using autocorrelation plots. Lawson proposed a Log-normal model being suitable in taking into account spatial correlation.

Let v_i measure the spatial random effects while τ_v^2 measure the variance while α be an overall level of the relative risk. The Log-normal model proposed by Lawson, et al. (2003) and applied by Johnson (2004) is given as (i=1,…, *n* are independent);

$$Y_i \sim Poisson(\theta E_i) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.16$$

$$\text{Log}\,\theta_i = \alpha + v_i \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.17$$

$$v_i \sim N(0, \tau_v^2) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.18$$

Probability function of the Poisson-log-normal distribution is not available analytically (Ntzoufras, 2009). Ntzoufras however, shows that the mean and variance are given by;

$$E(Y \mid \theta, \delta_\varepsilon^2) = \theta e^{\delta_\varepsilon^2/2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.19$$

and

$$V(Y \mid \theta, \delta_\varepsilon^2) = \theta e^{\delta_\varepsilon^2/2} + \theta^2 e^{2\delta_\varepsilon^2} - \theta^2 e^{\delta_\varepsilon^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.20$$

3.3.3 Besag, York and Mollie (BYM) Model

In this section we take into consideration a popular model introduced by Clayton and Kaldor (1987) and developed by Besag, York and Mollie after modifying the log-normal model described in equations 3.16 to 3.20. In this model for relative risk, area-specific random effects are decomposed into two components. One of the component accounts for the effects that vary in a structured manner in space (clustering or correlated heterogeneity). The second component models the effects that vary in an unstructured way between areas (uncorrelated heterogeneity). When undertaking CAR modeling, it is necessary to define an adjacency matrix that characterizes the neighborhood structure of the data being analyzed. There are several

approaches to doing this, including defining neighbors according to the distances between centroids and declaring two or more regions or districts to be neighbors if they share a boundary. A study by Best, Cockings, Bennett, Wakefield, and Elliott (2001) found out that the adjacency based neighborhood structure provided a better fit of the data than the distance approach, based on Deviance Information Criteria (DIC) comparisons. However, results by Arul, et al. (2007) found out the converse with the distance based matrix performing better than the adjacent based neighborhood structure. The use of centroid or distance approach appears not to offer a relative advantage over the other and may simply depend on the problem at hand. In this study, the neighborhood approach was used given the fact that data is not readily available on the centroid or distance matrix.

Let α be an overall level of the relative risk, u_i is the correlated heterogeneity and v_i is the uncorrelated heterogeneity.

Then under-five deaths Y_i, is assumed to follow Poisson distribution thus;

$$Y_i \sim Poisson(\theta E_i) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.21$$

$$\text{Log}\,\theta_i = \alpha + u_i + v_i \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.22$$

The distribution model for the uncorrelated heterogeneity is;

$$v_i \sim N(0, \tau_v^2) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.23$$

For the clustering or regional component, a spatial correlation structure is used, where estimation of the risk in any area depends on neighboring areas. The conditional autoregressive (CAR) model is thus given;

$$\left[u_i \mid u_j, i \neq j, \tau_u^2\right] \sim N(\bar{u}_i, \tau_i^2) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.24$$

Where

$$\bar{u}_i = \frac{1}{\sum_j w_{ij}} \sum_j u_j w_{ij}$$

$$\tau_i^2 = \frac{\tau_u^2}{\sum_j w_{ij}}$$

$w_{ij} = 1$ if i, j are adjacent (or 0 if they are not).

3.4 Priors used

Bolstand (2007) proposed three different priors for the Poisson distribution to include;

Positive uniform prior density, Jeffreys prior for Poisson and a gamma prior. Priors which are believed to have minimal impact on the posterior distribution are also often used and these are referred to as flat or noninformative or diffuse or vague prior.

3.4.1 Positive Uniform Prior Density

When there is no idea about the value of θ_i prior to looking at the data then equal weights can be assigned to the value of θ_i. Then a uniform prior density can be assigned i.e

$g(\theta_i)=1$ *for* $\theta_i>0$

……..……...3.25

In this case, the posterior will be $\propto 1 \,.\, \theta_i \sum_{i=1}^{n} y_i e^{-n\theta_i}$ ……...……..…...…..….3.26

3.4.2 Flat, noninformative, or diffuse or vague prior

A prior is noninformative if it has minimal impact on the posterior distribution. In the Poisson distribution, the flat, uninformative, or diffuse or vague prior is given as;

$\pi(\theta)=k=\frac{1}{b-a}$ *for* $a\leq\theta\leq b$ ……*3.27*

This conveys the fact that we have no a priori reason to favor any particular parameter value over another. With a flat prior, the posterior is just a constant multiplied with the likelihood as given below;

$g(\theta_i|y_i)=CL(\theta_i|y_i)$

……..3.28

where C is a constant and $L(\theta_i|y_i)$ is the likelihood function of θ_i given the data,*y*.

3.4.3 Jeffreys Prior

The Jeffreys prior for the Poisson is given as;

$g(\theta_i)\propto\frac{1}{\sqrt{\theta_i}}$ for $\theta_i>0$

……..3.29

Hence the posterior under Jeffereys prior will be;

$g(\theta_i|y_1,\ldots,y_n)\propto\frac{1}{\sqrt{\theta_i}}\theta_i\sum_{i=1}^{n}y_i e^{-n\theta_i}$ ……...3.30

Both the positive uniform prior density, flat, uninformative, or diffuse and the Jeffreys priors for the Poisson are considered as improper since their integral over all possible values is infinite.

3.4.4 Gamma Prior

Bolstand (2007) argues that when we have a single Poisson(θ) observation, the shape of the posterior using a gamma (α,β) is given by;

$$g(\theta_i \mid y_i) \propto g(\theta_i) f(y_i \mid \theta_i) \propto \frac{\beta^{\alpha} \theta_i^{\alpha-1} e^{-\beta\theta_i}}{\Gamma(\alpha)} \frac{\theta_i^{\sum_{i=1}^{n} y_i} e^{-\theta_i}}{\prod_{i=1}^{n} y_i!} \propto \theta_i^{\alpha-1+y} e^{-(\beta+1)\theta_i} \quad \text{.....3.31}$$

with the posterior mean and variance given respectively as;

$$E(\theta_i|y_i) = \frac{\alpha + y}{\beta + 1} \text{ and } \text{var}(\theta_i|y_i) = \frac{\alpha + y_i}{(\beta + 1)^2} \quad \text{.....3.32}$$

Use of the Bayesian statistics are based on the fact that there is uncertainty about the true value of the parameter and hence considers them as random variable with a given distribution. Lawson, et al. (2003) show that for a Poisson with a gamma prior, then

$y_i|\theta_i \sim$ Poisson $(E_i\theta_i)$ 3.33

$\theta_i|\alpha,\beta \sim$ Gamma(α,β).....3.34

$\alpha|\upsilon \sim h_{\alpha}(\upsilon)$

.....3.35

$\beta|v \sim h_{\beta}(\rho)$

.....3.36

The above shows a distribution in a form of a hierarchy but can be terminated at second level (equation 3.35 and 3.36) and is described as the hierarchical Bayesian approach.

3.5 HB Model Selection

Selection of the best model was assessed using the Deviance Information Criteria (DIC) supported by a number of studies (Earnest, et al., 2010; Ramis-Prieto, et al., 2007). DIC was introduced by Spiegelhalter, Best, Carlin, and Der-Linde, (2002) for model comparison. DIC is a hierarchical modelling generalization of the Akaike information criterion (AIC) and the Bayesian information criterion (BIC), also known as the Schwarz criterion. It is particularly useful in the Bayesian model selection problems where the posterior distributions of the models have been obtained by the Markov chain Monte Carlo (MCMC) simulation.

Define the deviance as D(θ) = -2log(p(y|θ))+C, where ,*y,* are the data, *θ,* are the unknown parameters of the model and p(y|θ) is the likelihood function. *C* is a constant that cancels out in all calculations that compare different models, and which therefore does not need to be known (Spiegelhalter, et al., 2002).

The expectation $\overline{D} = E^{\theta}[D(\theta)]$ is a measure of how well the model fits the data; the larger this is, the worse the fit.

The effective number of parameters of the model is computed as, $p_D = \overline{D} - D(\overline{\theta})$, where $\overline{\theta}$ is the expectation of *θ*. The deviance information criterion is calculated as $DIC = p_D + \overline{D}$ 3.37

The idea is that models with smaller DIC should be preferred to models with larger DIC (Chen, 2009). Roughly differences of two models with more than 10 values points will show that they

are substantially different (Holsinger 2012). Differences in DIC less than 5 could be misleading to report the model with lowest DIC. The advantage of DIC over other criteria in case of the Bayesian model selection is that the DIC is easily calculated from the samples generated by a MCMC simulation (Berg, Meyer, and Yu, 2004; Spiegelhalter, et al., 2002). AIC and BIC require calculating the likelihood at its maximum over θ_i, which is not readily available from the MCMC simulation.

3.6 Simulation

WinBUGS version 1.4 software as recommended by Congdon (2001); Lawson (2008); Lawson, et al. (2003); Rao (2003) in the Bayesian setting was used to monitor convergence and to estimate the various model parameters. Convergence was assessed in particular by use of dynamic trace, Kernel density plots and the Gelman-Rubin scale reduction factor which are included in the WinBUGS 1.4 software. Figure 3.2 demonstrates part of the code that was used in the WinBUGS 1.4 software. In this study, the main parameters that were estimated were relative disease risk and the variance in the various models.

```
model
{
for (i in 1:N){
                    y[i] ~ dpois(mu[i])
                    mu[i]<-e[i]*theta[i]
                    theta[i]~dgamma(a,b)
                    }

a~dexp(0.1)
b~dexp(0.1)

mean<-a/b
var<-a/pow(b,2)
}
```

Figure 3.2: Code fragment used in the WinBUGS Software for Poisson-gamma Model

3.7 Assumptions

In this study, the key assumptions made related to models developed were that the effects of neighboring countries are taken care of in the random error term. Additionally the within geographical boundaries defined by such features like lakes, rivers, mountains and forests are also taken care of in the random error term of the models used.

CHAPTER FOUR: SMALL AREA ESTIMATION: FOCUS ON UNDER FIVE MORTALITY DATA IN UGANDA

4.1 Comparison of SMR to Poisson-gamma, Log-normal and the BYM Using UDHS 1995 Data

The results presented in Table 4.1 show that the traditional method of estimation using SMR did not show variability from the model approach. This may largely be attributed to the fact that the numbers of districts were few by 1995 and the observed counts were fairly substantial. By 1995, there were a total of 38 districts although the demographic and health survey covered 37 due to insecurity in one of the districts. Compared to 2006 when the country had a total of 56 districts and even if the sample size had slightly increased, SMR showed more reliable and stable estimates with fewer districts for either 1995 or 2001.

The standard deviation (sd) under Log-normal and the BYM models were higher than those from the Poisson-gamma model. However, the DIC for Log-normal and the BYM were lower under the Poisson-gamma model but not substantial to warrant a difference in the two models (less than 10). It was further observed that convergence for both models (Poisson-gamma, log-normal and the BYM) were achieved at 4,000 iterations.

The results provided in Table 4.1 further show that the districts of Kalangala, Bundibugyo, Kamuli, Kiboga, Kotido, Mbarara and Kibaale had high relative risk of under-five mortality by 1995. Kumi district also exhibited a slightly high relative risk of under-five mortality by about four percent compared to the national average.

Table 4.1: Comparison of RR from SMR, Poisson-gamma, Log-normal and BYM models using UDHS 1995 Data

No.	District	under-five deaths (y_i)	SMR	Poisson-Gamma model (DIC=310.4)	Log-normal model (DIC=305.4)	BYM (DIC=305.9)
1	Kalangala	21	4.16	2.09	2.66	2.65
2	Bundibugyo	54	1.62	1.45	1.49	1.49
3	Kamuli	221	1.19	1.17	1.17	1.17
4	Kiboga	54	1.24	1.16	1.17	1.16
5	Kotido	29	1.22	1.10	1.12	1.11
6	Mbarara	295	1.11	1.10	1.10	1.10
7	Kibaale	70	1.11	1.07	1.07	1.07
8	Kumi	94	1.04	1.01	1.01	1.01
9	Mubende	145	0.94	0.93	0.93	0.93
10	Kisoro	49	0.92	0.90	0.89	0.89
11	Masaka	228	0.88	0.88	0.88	0.88
12	Rakai	83	0.70	0.70	0.70	0.70
13	Pallisa	96	0.70	0.70	0.70	0.70
14	Lira	145	0.70	0.70	0.70	0.70
15	Jinja	76	0.68	0.69	0.69	0.68
16	Arua	176	0.67	0.67	0.67	0.67
17	Hoima	37	0.66	0.67	0.66	0.66
18	Soroti	106	0.64	0.65	0.65	0.65
19	Tororo	136	0.64	0.64	0.64	0.64
20	Gulu	89	0.63	0.64	0.64	0.64
21	Bushenyi	133	0.63	0.64	0.64	0.63
22	Kabarole	134	0.63	0.63	0.63	0.63
23	Nebbi	81	0.62	0.62	0.62	0.62
24	Kampala	147	0.62	0.62	0.62	0.62
25	Luwero	85	0.61	0.62	0.62	0.62
26	Mbale	166	0.61	0.61	0.61	0.61
27	Iganga	217	0.60	0.60	0.60	0.60
28	Moyo	42	0.58	0.60	0.59	0.59
29	Mpigi	152	0.54	0.54	0.54	0.55
30	Kapchorwa	20	0.45	0.49	0.49	0.49
31	Moroto	33	0.46	0.49	0.49	0.49
32	Mukono	119	0.47	0.48	0.48	0.48
33	Apac	87	0.46	0.47	0.47	0.47
34	Rukungiri	49	0.44	0.46	0.46	0.46
35	Masindi	28	0.38	0.41	0.42	0.42
36	Kasese	34	0.35	0.38	0.38	0.38
37	Kabale	34	0.29	0.32	0.32	0.32

Coefficient of variation (CV) using SMR for the 1995 UDHS data show minimal variability or simply the standard deviations (sd) were small (Table 4.2). In all the cases for the 37 districts none of the CVs exceeded 60%. The results further show that when few districts are involved in estimation of SMR and when substantial data are provided, SMR appear to be a good estimator of relative risk of under-five mortality.

The results obtained using UDHS data of 1995 generally show that there was no added advantage in using the model approach compared to the use of SMR. The *CV* from the models (Poisson-gamma, log-normal and BYM) were low with none exceeding 30 percent.

Table 4.2: Variation arising from use of SMR as shown by Coefficient of Variation (CV percent) using 1995 UDHS data

No.	District	sd for SMR	CV for SMR (%)	sd for Poisson-Gamma	CV for Poisson-Gamma (%)	sd for Log-normal	CV for Log-normal (%)	sd for BYM	CV for BYM (%)
1	Kalangala	0.218	5.2	0.4718	22.7	0.6607	24.8	0.6881	25.9
2	Bundibugyo	0.136	8.4	0.1924	13.3	0.2049	13.7	0.2110	14.2
3	Kamuli	0.067	5.6	0.0790	6.7	0.0792	6.7	0.0843	7.2
4	Kiboga	0.136	11.0	0.1534	13.3	0.1551	13.3	0.1592	13.7
5	Kotido	0.186	15.2	0.1886	17.1	0.1982	17.7	0.2037	18.3
6	Mbarara	0.058	5.2	0.0634	5.8	0.0659	6.0	0.0686	6.2
7	Kibaale	0.120	10.8	0.1228	11.5	0.1252	11.7	0.1296	12.1
8	Kumi	0.103	9.9	0.1016	10.0	0.1035	10.2	0.1031	10.2
9	Mubende	0.083	8.8	0.0764	8.2	0.0774	8.3	0.0774	8.4
10	Kisoro	0.143	15.5	0.1201	13.4	0.1235	13.8	0.1236	13.9
11	Masaka	0.066	7.5	0.0568	6.5	0.0573	6.6	0.0590	6.7
12	Rakai	0.110	15.7	0.0747	10.6	0.0753	10.7	0.0749	10.7
13	Pallisa	0.102	14.6	0.0703	10.0	0.0690	9.8	0.0692	9.9
14	Lira	0.083	11.9	0.0574	8.2	0.0566	8.1	0.0579	8.3
15	Jinja	0.115	16.8	0.0751	10.9	0.0754	11.0	0.0758	11.1
16	Arua	0.075	11.3	0.0504	15.4	0.0495	7.4	0.0489	7.3
17	Hoima	0.164	25.0	0.1028	7.5	0.0999	15.1	0.1023	15.4
18	Soroti	0.097	15.1	0.0616	9.5	0.0615	9.5	0.0611	9.5
19	Tororo	0.086	13.5	0.0543	8.5	0.0544	8.5	0.0542	8.5
20	Gulu	0.106	16.7	0.0662	10.4	0.0652	10.2	0.0652	10.2

No.	District	sd for SMR	CV for SMR (%)	sd for Poisson-Gamma	CV for Poisson-Gamma (%)	sd for Log-normal	CV for Log-normal (%)	sd for BYM	CV for BYM (%)
21	Bushenyi	0.087	13.7	0.0535	8.3	0.0540	8.5	0.0533	8.4
22	Kabarole	0.086	13.7	0.0539	8.5	0.0528	8.4	0.0532	8.4
23	Nebbi	0.035	5.7	0.0668	10.7	0.0675	10.9	0.0654	10.5
24	Kampala	0.082	13.3	0.0647	10.5	0.0498	8.0	0.0509	8.2
25	Luwero	0.108	17.6	0.0501	8.0	0.0642	10.4	0.0639	10.3
26	Mbale	0.078	12.8	0.0471	7.6	0.0465	7.6	0.0471	7.7
27	Iganga	0.068	11.4	0.0411	6.8	0.0399	6.6	0.0403	6.7
28	Moyo	0.154	26.7	0.0864	14.6	0.0847	14.3	0.0844	14.3
29	Mpigi	0.081	15.0	0.0431	8.0	0.0433	8.0	0.0444	8.1
30	Kapchorwa	0.224	50.1	0.0969	19.7	0.0940	19.1	0.0961	19.5
31	Moroto	0.174	38.1	0.0772	15.9	0.0773	15.8	0.0788	16.2
32	Mukono	0.092	19.7	0.0431	9.0	0.0423	8.9	0.0441	9.2
33	Apac	0.107	23.2	0.0494	10.4	0.0491	10.4	0.0505	10.7
34	Rukungiri	0.143	32.5	0.0614	13.2	0.0609	13.2	0.0635	13.7
35	Masindi	0.189	50.2	0.0703	17.0	0.0691	16.5	0.0714	17.1
36	Kasese	0.171	49.3	0.0596	15.8	0.0594	15.5	0.0630	16.4
37	Kabale	0.171	59.8	0.0496	15.7	0.0511	15.8	0.0535	16.5

Table 4.2 Continued

4.2 Comparison of SMR to Poisson-gamma, the BYM and Log-normal Models Using UDHS 2001 Data

Results presented in Table 4.3 show that traditional method of estimation using SMR were not different by much from the model approach when UDHS 2001 data was used. Again this may largely be attributed to the fact that the numbers of districts were still few by 2001 and the observed counts (y_i) were fairly substantial. In year 2001 only 34 districts were covered implying that higher deaths counts (y_i) per district were obtained compared to using UDHS 2006 data.

In about 56 percent and 74 percent of the cases, the standard deviation (sd) under Log-normal and the BYM models respectively were higher than those from the Poisson-gamma model. Results on the DIC appear to reveal no difference between Log-normal, BYM and the Poisson-gamma models (difference less than 5) though the latter had the least DIC value. Convergence for both models (Poisson-gamnma, log-normal and the BYM) were achieved at 4,000 iterations.

Results show that the district of Kalangala consistently had a very high SMR and RR of under-five mortality for the period 1995 to 2001. The other districts with high RR included; Rukungiri, Kisoro, Kabarole. The districts of Mpigi, Moyo and Mubende also had a slightly high relative risk of under-five mortality. Overall, results obtained using the four methods (SMR, Poisson-gamma, log-normal and the BYM model) were considerably not deviating from each other by much using the 1995 and 2001 datasets. This may be attributed to the fact that the numbers of districts were still few (34) and subsequently a large number of observations per district to reduce the noise. Taking the

rule of thumb for where $y_i<30$ (small sample size), we have only two districts of Kalangala and Masindi that break the rule with few observations (Table 4.3).

Table 4.3: Comparison of RR from SMR, Poisson-gamma, Log-normal and the BYM Models using 2001 data

No.	District	under-five deaths (y_i)	SMR	Poisson-gamma model (DIC=279.8)	Log-normal model (DIC=280.9)	BYM (DIC=281.1)
1	Kalangala	18	2.06	1.62	1.77	1.76
2	Rukungiri	152	1.68	1.64	1.65	1.66
3	Kisoro	113	1.56	1.52	1.54	1.54
4	Kabarole	152	1.30	1.28	1.28	1.28
5	Mpigi	111	1.08	1.07	1.07	1.07
6	Moyo	67	1.04	1.02	1.02	1.02
7	Mubende	177	1.02	1.02	1.02	1.02
8	Iganga	194	1.00	0.99	0.99	0.99
9	Kamuli	191	0.99	0.98	0.98	0.98
10	Masaka	189	0.98	0.97	0.97	0.97
11	Kiboga	52	0.90	0.89	0.89	0.87
12	Moroto	50	0.80	0.79	0.78	0.77
13	Rakai	90	0.76	0.76	0.76	0.76
14	Mbale	145	0.74	0.74	0.73	0.73
15	Soroti	71	0.70	0.70	0.70	0.70
16	Nebbi	95	0.66	0.66	0.66	0.66
17	Kumi	69	0.65	0.65	0.64	0.64
18	Bushenyi	151	0.63	0.63	0.63	0.63
19	Tororo	91	0.62	0.62	0.62	0.62
20	Mbarara	221	0.62	0.62	0.62	0.62
21	Apac	134	0.59	0.59	0.59	0.59
22	Kabale	85	0.57	0.57	0.57	0.57
23	Arua	156	0.57	0.57	0.57	0.57
24	Mukono	110	0.55	0.55	0.55	0.55
25	Kibaale	71	0.53	0.54	0.54	0.54
26	Pallisa	72	0.50	0.51	0.51	0.51
27	Jinja	53	0.50	0.51	0.50	0.51
28	Luwero	59	0.49	0.50	0.50	0.50
29	Lira	120	0.49	0.49	0.49	0.49
30	Kampala	106	0.35	0.36	0.36	0.36
31	Masindi	30	0.20	0.21	0.22	0.22
32	Hoima	19	0.17	0.19	0.20	0.19
33	Kapchorwa	3	0.06	0.11	0.13	0.13
34	Kotido	22	0.11	0.13	0.13	0.13

NB: Districts not covered in the survey due to insecurity: Kasese, Bundibugyo, Gulu and Kitgum

Districts that had higher *CVs* had peculiar SMR that were quite of out-of-range. For instance, Kapchorwa shows an SMR of 0.06 implying a very low chance of relative risk to under-five mortality in the district to almost 94% below the national level. However, the three models (Posson-gamma, log-normal and BYM) independently show close and comparable results and their *CV*s were considerably low (<100% in all cases). These findings together with test for reliability of SMR using *CV* show that there are higher possibilities of "noise" in using SMR. Results obtained using the model approaches are also in close range with those from neighbouring districts especially for districts where SMR appears to be out-of-range.

SMR results using the 2001 UDHS data showed high variability (CV>100%) in only three districts of Kapchorwa, Kotido and Hoima as shown in Table 4.4. Overall, other district's coefficients of variation were relatively low indicating low level of 'noise' in the SMR computations using 1995 and 2001 datasets.

Table 4.4: Variation arising from use of SMR as shown by Coefficient of Variation (CV percent) using 2001 UDHS data

No.	District	sd for SMR	CV for SMR %	sd for Poisson-gamma	CV for Poisson-gamma %	sd for Log-normal	CV for Log-normal %	sd for BYM	CV for BYM %
1	Kalangala	0.24	11.45	0.3615	22.2	0.4205	23.8	0.4308	24.5
2	Rukungiri	0.08	4.81	0.1326	8.1	0.1370	8.3	0.1395	8.5
3	Kisoro	0.09	6.01	0.1433	9.4	0.1479	9.6	0.1466	9.6
4	Kabarole	0.08	6.24	0.1021	8.0	0.1047	8.2	0.1082	8.4
5	Mpigi	0.10	8.77	0.1009	9.4	0.0997	9.3	0.1031	9.7
6	Moyo	0.12	11.74	0.1201	11.8	0.1243	12.2	0.1250	12.2
7	Mubende	0.08	7.34	0.0757	7.4	0.0766	7.5	0.0783	7.7
8	Iganga	0.07	7.20	0.0712	7.2	0.0710	7.2	0.0726	7.3
9	Kamuli	0.07	7.30	0.0714	7.3	0.0712	7.3	0.0727	7.4
10	Masaka	0.07	7.48	0.0678	7.0	0.0712	7.3	0.0703	7.2
11	Kiboga	0.14	15.41	0.1186	13.4	0.1195	13.5	0.1234	14.0
12	Moroto	0.14	17.73	0.1089	13.8	0.1090	13.9	0.1113	14.3
13	Rakai	0.11	13.79	0.0793	10.4	0.0786	10.4	0.0784	10.3

No.	District	sd for SMR	CV for SMR %	sd for Poisson-gamma	CV for Poisson-gamma %	sd for Log-normal	CV for Log-normal %	sd for BYM	CV for BYM %
14	Mbale	0.08	11.26	0.0611	8.3	0.0596	8.1	0.0603	8.3
15	Soroti	0.12	16.98	0.0812	11.5	0.0810	11.6	0.0814	11.7
16	Nebbi	0.10	15.62	0.0675	10.3	0.0659	10.0	0.0666	10.2
17	Kumi	0.12	18.56	0.0753	11.6	0.0763	11.8	0.0752	11.6
18	Bushenyi	0.08	12.86	0.0505	8.0	0.0509	8.1	0.0506	8.0
19	Tororo	0.11	16.97	0.0643	10.3	0.0644	10.4	0.0638	10.3
20	Mbarara	0.07	10.82	0.0419	6.7	0.0414	6.7	0.0414	6.7
21	Apac	0.09	14.53	0.0510	8.6	0.0506	8.5	0.0511	8.6
22	Kabale	0.11	19.09	0.0602	10.6	0.0597	10.5	0.0608	10.7
23	Arua	0.08	14.16	0.0446	7.9	0.0450	7.9	0.0455	8.0
24	Mukono	0.10	17.26	0.0527	9.5	0.0514	9.3	0.0521	9.4
25	Kibaale	0.12	22.30	0.0628	11.6	0.0621	11.6	0.0638	11.9
26	Pallisa	0.12	23.37	0.0610	11.9	0.0583	11.4	0.0585	11.4
27	Jinja	0.14	27.44	0.0672	13.3	0.0668	13.2	0.0679	13.5
28	Luwero	0.13	26.49	0.0628	12.6	0.0626	12.6	0.0625	12.6
29	Lira	0.09	18.61	0.0440	8.9	0.0446	9.1	0.0449	9.1
30	Kampala	0.10	27.34	0.0352	9.8	0.0353	9.8	0.0363	10.0
31	Masindi	0.18	91.89	0.0376	17.6	0.0380	17.6	0.0425	19.4
32	Hoima	0.23	135.77	0.0406	21.4	0.0415	21.1	0.0460	23.5
33	Kapchorwa	0.58	1003.02	0.0455	41.2	0.0484	36.2	0.0530	39.9
34	Kotido	0.21	189.68	0.0257	20.4	0.0274	20.7	0.0356	26.7

4.3 Comparison of SMR to Poisson-gamma, BYM and Log-normal Models Using UDHS 2006 Data

Results after using SMR, Poisson-gamma, Log-normal and the BYM, show that there was a lot of "noise" from the SMR compared to model approach using the UDHS 2006 data. The "noise" from SMR results can be attributed to the fact that more districts (56) were introduced by year 2006 that reduced the sample size per district. DIC results suggest that all the three models; Poisson-gamma, Log-normal and the BYM models had no substantial difference amongst them. It was also observed from the diagnostics test for the presence of autocorrelation that its existence was rather feeble. The absence of a stronger autocorrelation suggests that the Poisson-gamma model could as well perform as

good as the other model options. Additionally, results presuppose that either the Poisson-gamma or the BYM or the log-normal appear to be suitable for modeling and estimation of relative risk of under-five in Uganda.

Findings also show that the Poisson-gamma model potentially perform better than other HB models as shown in Figure 4.1 below. The model with the lowest DIC is considered better. It is also important to note that the differences between the values of the DIC were generally low indicating that all the three models were fairly good in deriving relative risk of under-five mortality.

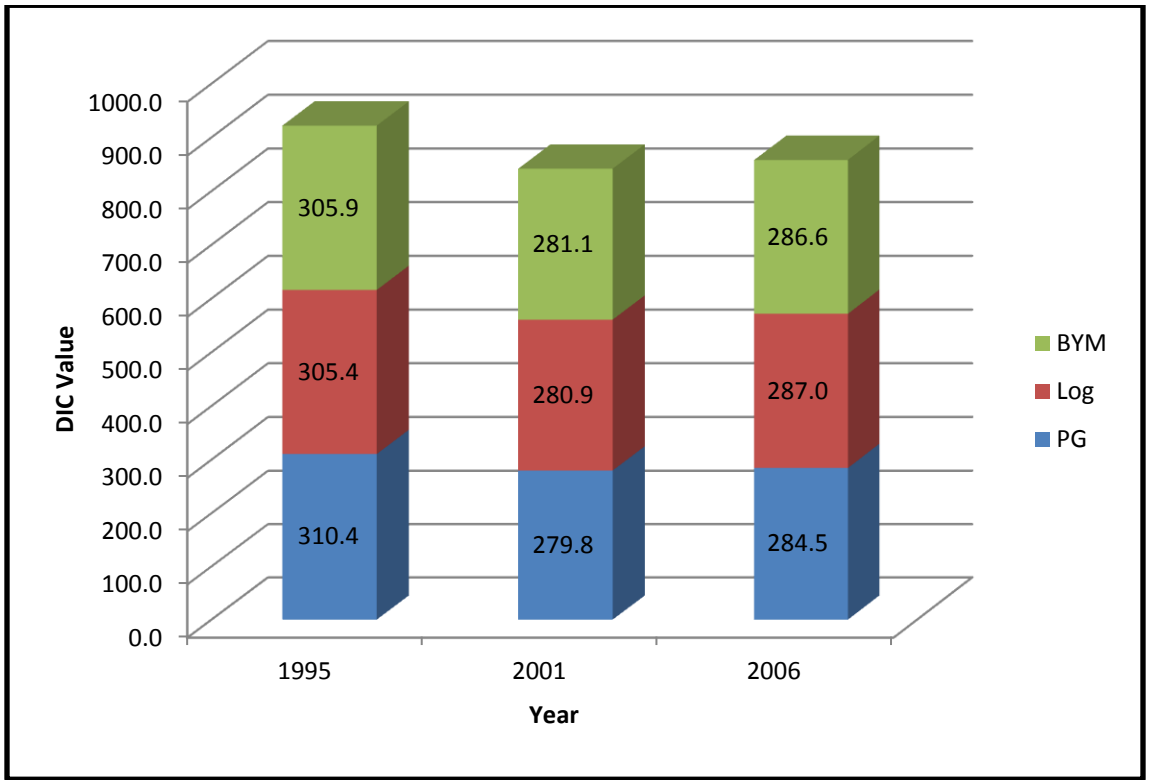


Figure 4.1: Comparison of Model Performance using DIC

Using 2006 data, the districts of Moroto, Kotido, Gulu, Iganga, Kamuli, Pader, Mbale, Mubende and Arua were identified with high relative risk (>=1.10) of under-five mortality. Sembabule, Kitgum, Nakapiripirit, Kisoro and Kamwenge had equally high relative risk of under-five mortality (Table 4.5).

Table 4.5: Comparison of SMR to Poisson-gamma, BYM and log-normal models using 2006 UDHS data

No.	District	under-five deaths (y_i)	SMR	RR - Poisson(DIC =284.5)	RR -BYM (DIC=286.6)	RR - Lognormal (DIC= 287.0)
1	Moroto	111	1.53	1.28	1.46	1.31
2	Kotido	241	1.06	1.27	1.36	1.27
3	Gulu	210	1.00	1.21	1.26	1.20
4	Iganga	192	0.96	1.17	1.18	1.16
5	Kamuli	192	0.96	1.16	1.20	1.16
6	Pader	139	0.97	1.13	1.20	1.11
7	Mbale	167	0.91	1.12	1.13	1.11
8	Mubende	177	0.90	1.12	1.12	1.10
9	Arua	273	0.84	1.11	1.12	1.10
10	Sembabule	66	1.05	1.05	1.05	1.04
11	Kitgum	108	0.87	1.05	1.12	1.03
12	Nakapiripirit	61	1.03	1.03	1.08	1.03
13	Kisoro	8	0.09	1.03	1.03	1.02
14	Kamwenge	74	0.88	1.02	0.97	1.00
15	Bundibugyo	57	0.85	0.98	0.94	0.96
16	Rakai	111	0.67	0.93	0.90	0.92
17	Nakasongola	77	0.64	0.87	0.86	0.87
18	Kibaale	174	0.59	0.87	0.84	0.86
19	Kalangala	23	0.58	0.87	0.80	0.85
20	Apac	78	0.58	0.85	0.86	0.85
21	Tororo	68	0.57	0.85	0.84	0.85

No.	District	under-five deaths (y_i)	SMR	RR - Poisson(DIC =284.5)	RR -BYM (DIC=286.6)	RR - Lognormal (DIC= 287.0)
22	Busia	7	0.55	0.85	0.82	0.84
23	Kyenjojo	35	0.57	0.84	0.82	0.84
24	Masindi	79	0.54	0.83	0.82	0.82
25	Hoima	57	0.52	0.83	0.78	0.82
26	Kiboga	156	0.49	0.82	0.80	0.81
27	Bushenyi	32	0.54	0.82	0.78	0.81
28	Kampala	87	0.52	0.82	0.42	0.81
29	Nebbi	127	0.51	0.81	0.80	0.80
30	Masaka	140	0.52	0.81	0.78	0.80
31	Kapchorwa	65	0.39	0.79	0.79	0.80
32	Pallisa	19	0.49	0.79	0.79	0.79
33	Kaberamaido	52	0.27	0.78	0.77	0.78
34	Bugiri	9	0.45	0.78	0.74	0.78
35	Sironko	104	0.36	0.76	0.76	0.76
36	Mukono	26	0.46	0.75	0.71	0.76
37	Katakwi	29	0.35	0.75	0.75	0.76
38	Kanungu	27	0.36	0.75	0.70	0.75
39	Soroti	147	0.36	0.74	0.73	0.74
40	Kayunga	34	0.32	0.74	0.70	0.74
41	Lira	27	0.45	0.73	0.75	0.74
42	Kasese	66	0.40	0.73	0.69	0.73
43	Kumi	34	0.34	0.73	0.74	0.73

No.	District	under-five deaths (y_i)	SMR	RR - Poisson(DIC =284.5)	RR -BYM (DIC=286.6)	RR - Lognormal (DIC= 287.0)
44	Rukungiri	49	0.33	0.71	0.67	0.73
45	Luwero	36	0.36	0.71	0.68	0.72
46	Ntungamo	52	0.34	0.69	0.65	0.70
47	Kabale	65	0.36	0.69	0.65	0.70
48	Moyo	177	0.17	0.68	0.63	0.69
49	Mbarara	13	0.41	0.66	0.66	0.69
50	Jinja	41	0.24	0.66	0.63	0.68
51	Mpigi	26	0.29	0.66	0.62	0.68
52	Kabarole	26	0.23	0.65	0.62	0.67
53	Adjumani	10	0.13	0.64	0.63	0.67
54	Yumbe	16	0.16	0.63	0.60	0.66
55	Mayuge	8	0.09	0.60	0.49	0.64
56	Wakiso	63	0.20	0.49	0.55	0.54

Table 4.5 (Continued)

Using 2006 dataset again, results show high *CV*s for the districts of Adjumani, Kaberamaido, Kisoro, Mayuge, Moyo and Yumbe. The *CV*s for these districts using the SMR were relatively high (>100%) as shown in Table 4.6 depicting unreliability in utilization of SMR to estimate relative risk of under-five mortality. The *CV*s for the Poisson-gamma model were all much lower than 50% compared to those from SMR (Table 4.6).

Table 4.6: Variability arising from use of SMR as shown by Coefficient of Variation (CV percent) using 2006 UDHS data

No.	District	sd for SMR	*CV* for SMR %	sd-Poisson	*CV* for Poisson-gamma %	sd-Lognormal	*CV* for Log-normal %	sd-BYM	*CV* for BYM %
1	Moroto	0.095	6.2	0.30650	23.4	0.36970	28.2	0.43030	29.5
2	Kotido	0.064	6.0	0.22710	17.9	0.25240	19.9	0.26900	19.8
3	Gulu	0.069	6.9	0.22620	18.9	0.24450	20.4	0.25640	20.3
4	Iganga	0.072	7.5	0.22820	19.7	0.24390	21.0	0.25270	21.4
5	Kamuli	0.072	7.5	0.22240	19.2	0.23750	20.5	0.26370	22.0
6	Pader	0.085	8.8	0.24200	21.8	0.25040	22.6	0.28680	23.9
7	Mbale	0.077	8.5	0.22240	20.0	0.23620	21.3	0.24190	21.4
8	Mubende	0.075	8.3	0.21810	19.8	0.19530	17.8	0.24130	21.5
9	Arua	0.061	7.3	0.18730	17.0	0.23120	21.0	0.20420	18.2
10	Sembabule	0.123	11.7	0.26530	25.5	0.28260	27.2	0.31230	29.7
11	Kitgum	0.096	11.0	0.24120	23.4	0.24400	23.7	0.28950	25.8
12	Nakapiripirit	0.128	12.4	0.27410	26.6	0.28110	27.3	0.33410	30.9
13	Kisoro	0.354	388.1	0.25920	25.4	0.26260	25.7	0.30120	29.2
14	Kamwenge	0.116	13.2	0.25650	25.7	0.26270	26.3	0.27680	28.5
15	Bundibugyo	0.132	15.5	0.26140	27.2	0.26100	27.2	0.27240	29.0
16	Rakai	0.095	14.1	0.20980	22.8	0.20420	22.2	0.21480	23.9

No.	District	sd for SMR	*CV* for SMR %	sd-Poisson	*CV* for Poisson-gamma %	sd-Lognormal	*CV* for Log-normal %	sd-BYM	*CV* for BYM %
17	Nakasongola	0.209	32.8	0.20990	24.1	0.25050	28.8	0.29590	34.4
18	Kibaale	0.114	19.2	0.16750	19.5	0.20840	24.2	0.22190	26.4
19	Kalangala	0.378	65.7	0.26080	30.7	0.27330	32.2	0.30040	37.6
20	Apac	0.076	13.1	0.20830	24.5	0.15490	18.2	0.17400	20.2
21	Tororo	0.113	19.8	0.21090	24.8	0.20000	23.5	0.21510	25.6
22	Busia	0.169	30.6	0.27410	32.6	0.22330	26.6	0.26130	31.9
23	Kyenjojo	0.121	21.4	0.24030	28.6	0.20680	24.6	0.21070	25.7
24	Masindi	0.113	21.0	0.20380	24.9	0.19690	24.0	0.20640	25.2
25	Hoima	0.132	25.4	0.21470	26.2	0.20020	24.4	0.21230	27.2
26	Kiboga	0.177	36.0	0.16320	20.1	0.22270	27.5	0.24340	30.4
27	Bushenyi	0.080	14.9	0.23620	29.2	0.15530	19.2	0.16190	20.8
28	Kampala	0.089	17.2	0.19300	23.8	0.16410	20.3	0.12620	30.0
29	Nebbi	0.107	20.8	0.17140	21.4	0.17950	22.4	0.20660	25.8
30	Masaka	0.085	16.4	0.16680	20.9	0.15740	19.7	0.17580	22.5
31	Kapchorwa	0.229	58.6	0.20150	25.2	0.22430	28.0	0.26040	33.0
32	Pallisa	0.124	25.3	0.24100	30.5	0.19180	24.3	0.20210	25.6
33	Kaberamaido	0.333	124.3	0.20790	26.7	0.23700	30.4	0.24250	31.5
34	Bugiri	0.139	31.0	0.24380	31.3	0.19520	25.0	0.20550	27.8
35	Sironko	0.196	54.5	0.16440	21.6	0.20920	27.5	0.23590	31.0
36	Mukono	0.098	21.3	0.21920	28.8	0.16810	22.1	0.17060	24.0
37	Katakwi	0.192	54.2	0.21790	28.7	0.20720	27.3	0.21440	28.6
38	Kanungu	0.186	52.3	0.22140	29.5	0.20190	26.9	0.22380	32.0
39	Soroti	0.171	47.5	0.15300	20.7	0.19600	26.5	0.21830	29.9
40	Kayunga	0.192	59.4	0.20980	28.4	0.19690	26.6	0.20770	29.7
41	Lira	0.082	18.2	0.21520	29.1	0.14290	19.3	0.14850	19.8
42	Kasese	0.123	31.1	0.17960	24.6	0.17380	23.8	0.17700	25.7

No.	District	sd for SMR	*CV* for SMR %	sd-Poisson	*CV* for Poisson-gamma %	sd-Lognormal	*CV* for Log-normal %	sd-BYM	*CV* for BYM %
43	Kumi	0.171	50.0	0.20400	27.9	0.18960	26.0	0.22020	29.8
44	Rukungiri	0.165	50.2	0.19110	26.2	0.18660	25.6	0.19930	29.7
45	Luwero	0.143	39.6	0.20170	28.0	0.17620	24.5	0.17520	25.8
46	Ntungamo	0.139	40.4	0.18170	26.0	0.17190	24.6	0.17130	26.4
47	Kabale	0.124	34.8	0.17540	25.1	0.16310	23.3	0.15930	24.5
48	Moyo	0.277	161.6	0.13040	18.9	0.19250	27.9	0.20530	32.6
49	Mbarara	0.075	18.4	0.20720	30.0	0.12660	18.3	0.12340	18.7
50	Jinja	0.196	82.3	0.18010	26.5	0.18380	27.0	0.18630	29.6
51	Mpigi	0.156	54.3	0.19000	27.9	0.17230	25.3	0.17070	27.5
52	Kabarole	0.196	85.8	0.19220	28.7	0.17650	26.3	0.17750	28.6
53	Adjumani	0.316	248.9	0.20670	30.9	0.18710	27.9	0.20600	32.7
54	Yumbe	0.250	153.2	0.19210	29.1	0.17660	26.8	0.19550	32.6
55	Mayuge	0.354	404.5	0.13090	20.5	0.18200	28.4	0.12460	25.4
56	Wakiso	0.126	63.5	0.19290	35.7	0.13200	24.4	0.17770	32.3

Table 4.6 Continued

4.4 Spatial Variations for the Relative Risk of Under-five Mortality in Uganda

Results presented in Table 4.7 show a consistent high *RR* of under-five mortality for the districts of Kotido, Kamuli, Mubende and Kisoro over the period 1995 to 2006. The districts of Jinja, Kabale, Luwero, Lira, Kasese Soroti, Mukono, Pallisa, Kapchorwa have relatively had low *RR* over the time period (1995 to 2006). Results also show some peculiar and rather a very low *RR* for the district of Kotido for the year 2001 and this could be associated with data problems despite the smoothing by the HB model approach. The results for Kotido deviates from the trend given the fact that in 1995 the district had a *RR* of 1.11 while in 2006 the figure was 1.21 compared to 0.13 in 2001.

Table 4.7: Variability of Relative Risk of Under-five over the period 1995 to 2006

No.	District	*RR* (2006)	*RR* (2001)	*RR* (1995)	No.	District	*RR* (2006)	*RR* (2001)	*RR* (1995)
1	Moroto	1.28	0.79	0.49	29	Masaka	0.81	0.97	0.88
2	Kotido	1.27	0.13	1.11	30	Nebbi	0.81	0.66	0.62
3	Gulu	1.21	n/a	0.64	31	Kapchorwa	0.79	0.11	0.49
4	Iganga	1.17	0.99	0.60	32	Pallisa	0.79	0.51	0.7
5	Kamuli	1.16	0.98	1.17	33	Bugiri	0.78	n/a	n/a
6	Pader	1.13	n/a	n/a	34	Kaberamaido	0.78	n/a	n/a
7	Mbale	1.12	0.74	0.61	35	Sironko	0.76	n/a	n/a
8	Mubende	1.12	1.02	0.93	36	Kanungu	0.75	n/a	n/a
9	Arua	1.11	0.57	0.67	37	Katakwi	0.75	n/a	n/a
10	Kitgum	1.05	n/a	n/a	38	Mukono	0.75	0.55	0.48
11	Sembabule	1.05	n/a	n/a	39	Kayunga	0.74	n/a	n/a
12	Kisoro	1.03	1.52	0.90	40	Soroti	0.74	0.7	0.65
13	Nakapiripirit	1.03	n/a	n/a	41	Kasese	0.73	n/a	0.38
14	Kamwenge	1.02	n/a	n/a	42	Kumi	0.73	0.65	1.01
15	Bundibugyo	0.98	n/a	1.45	43	Lira	0.73	0.49	0.70
16	Rakai	0.93	0.76	0.71	44	Luwero	0.71	0.5	0.62
17	Kalangala	0.87	1.62	2.10	45	Rukungiri	0.71	1.64	0.46
18	Kibaale	0.87	0.54	1.07	46	Kabale	0.69	0.57	0.32
19	Nakasongola	0.87	n/a	n/a	47	Ntungamo	0.69	n/a	n/a
20	Apac	0.85	0.59	0.47	48	Moyo	0.68	1.02	0.59
21	Busia	0.85	n/a	n/a	49	Jinja	0.66	0.51	0.69
22	Tororo	0.85	0.62	0.64	50	Mbarara	0.66	0.62	1.10

No.	District	*RR (2006)*	*RR (2001)*	*RR (1995)*	No.	District	*RR (2006)*	*RR (2001)*	*RR (1995)*
23	Kyenjojo	0.84	n/a	n/a	51	Mpigi	0.66	1.07	0.55
24	Hoima	0.83	0.19	0.67	52	Kabarole	0.65	1.28	0.63
25	Masindi	0.83	0.21	0.41	53	Adjumani	0.64	n/a	n/a
26	Bushenyi	0.82	0.63	0.64	54	Yumbe	0.63	n/a	n/a
27	Kampala	0.82	0.36	0.62	55	Mayuge	0.6	n/a	n/a
28	Kiboga	0.82	0.89	1.16	56	Wakiso	0.49	n/a	n/a

n/a = *Not Applicable (District was not created by then or was not surveyed due to insecurity)*

Figure 4.2 shows a spatial trend in the northern region for districts of Arua, Gulu, Pader, Kotido and Moroto. This trend could be associated with effects of war that ravaged the northern region for two decades since 1986 up to 2006 and the harsh socio-economic conditions and cattle rustling in the districts of Kotido and Moroto that could have lead to more deaths especially among the children under-five years of age.

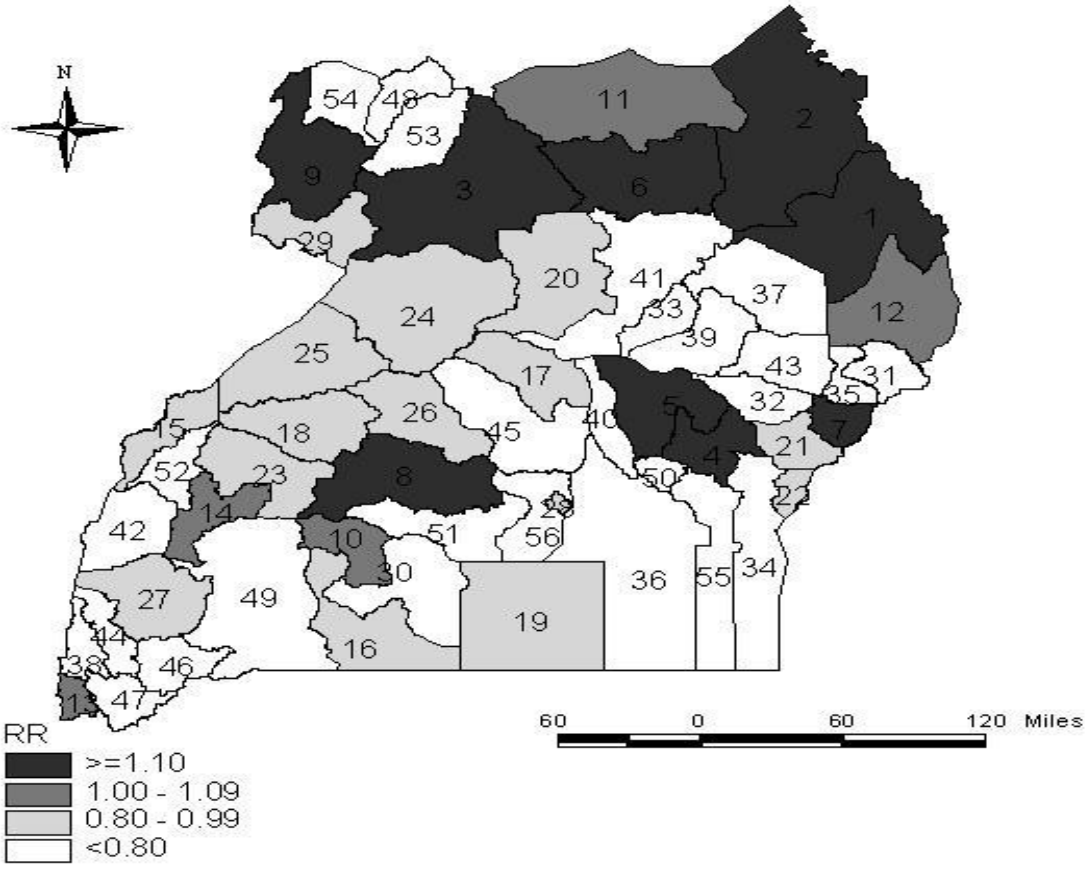


Figure 4.2: Spatial distribution showing under-five relative risk by 56 districts of Uganda (See numbers 1 - 56 in Table 4.6 above for districts)

4.5 Diagnostic Tests

Diagnostic tools were used to test for convergence and the suitability of models. This section provides the different diagnostic tests that were used.

4.5.1 Convergence Tests

Dynamic traces were used to test whether convergence is attained using graphical means (see Appendix 1). In this study, WinBUGS software was used to check for convergence. Accuracy of the estimates was also assessed by using the Monte Carlo standard errors of the mean. To reach efficiency, the Monte Carlo errors must be small in relation to the standard deviation (Lawson, et al., 2003). Using the Log-normal model and the UDHS2006 data, convergence was observed to have fairly been attained at 4000 iterations. If the model has converged, the trace plot will move "snake" around the mode of the distribution. Details of the sample of the first 20 out of 56 dynamic trace plots are shown in Appendix 1.

Further diagnostic tests were done using kernel density plots and results also revealed stability of the estimates using the Log-normal model at 4000 iteration. A more satisfactory kernel density plot would look more of a bell-shaped, though it need not be symmetric. A sample of the first 20 out of 56 kernel density plots are also shown in Appendix 2.

As already mentioned, geographical locations like districts treated as small area may exhibit some form of spatial autocorrelation. Spatial clustering of disease is almost inevitable since human populations generally live in spatial clusters rather than random distribution of space (Kleinschmidta, et al., 2000). Existence of spatial autocorrelation may therefore violate basic assumption used in the Poisson-gamma or Log-normal models. It was therefore imperative in this study to test for the existence of

autocorrelation and if it exists, a model taking care of its existence can be used. In a typical case or undesirable form, the autocorrelation plot will be a solid bar across the screen as shown in Figure 4.3.

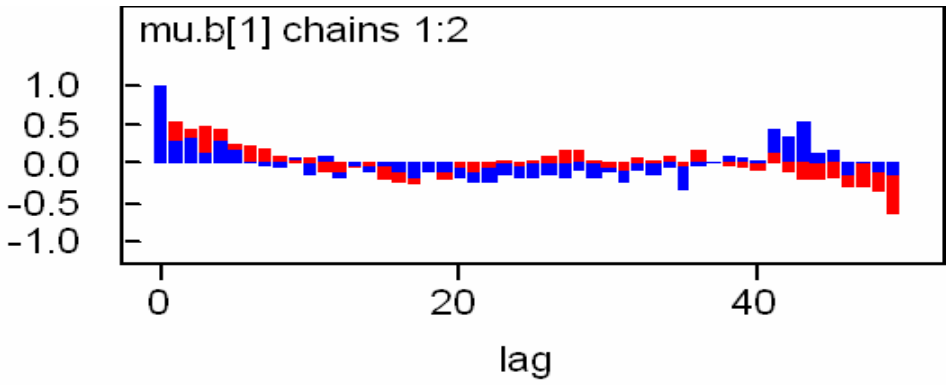


Figure 4.3: Graphical Example of Model with Autocorrelation (Spiegelhalter, Thomas, Best, and Lunn, 2003)

Results show that existence of autocorrelation was not evident as shown in Appendix 3.

4.5.2 Sensitivity Analysis

When prior information is available, sensitivity analysis focuses on the structure of the prior distribution; it focuses on how different choices of prior parameters may influence the posterior inference (Ntzoufras, 2009). Sensitivity analysis was conducted to investigate whether results in the models remain stable when different prior information is used. In this study, a diffuse or flat prior was used with the Poisson model (Appendix 4). Overall results show that after use of the flat and the gamma prior, the results show minimal difference in DIC. The main differences were that convergence using the flat prior were reached at a higher value (15,000 iterations) compared to use of the gamma prior (4,000 iterations). The other difference too is that the standard deviations (*sd*) using

the flat prior were slightly higher compared to use of the gamma prior as show in Table 4.8.

Table 4.8: Comparison of results from a Poisson-gamma model using gamma and flat priors using UDHS2006 data

No.	District	RR-Poisson with gamma prior (DIC=284.461)	RR Poisson with flat prior(DIC =286.274)	sd-gamma prior	sd – flat prior	No.	District	RR-Poisson with gamma prior(DIC= 284.461)	RR Poisson with flat prior(DIC= 286.274)	sd-gamma prior	sd – flat prior
1	Moroto	1.31	1.62	0.30650	0.4339	29	Nebbi	0.80	0.80	0.17140	0.2243
2	Kotido	1.27	1.43	0.22710	0.2724	30	Masaka	0.80	0.80	0.16680	0.1892
3	Gulu	1.20	1.36	0.22620	0.2761	31	Kapchorwa	0.80	0.75	0.20150	0.3147
4	Iganga	1.16	1.30	0.22820	0.2736	32	Pallisa	0.79	0.78	0.24100	0.2418
5	Kamuli	1.16	1.30	0.22240	0.2748	33	Kaberamaido	0.78	0.71	0.20790	0.3327
6	Pader	1.11	1.27	0.24200	0.3046	34	Bugiri	0.78	0.74	0.24380	0.2476
7	Mbale	1.11	1.24	0.22240	0.276	35	Sironko	0.76	0.70	0.16440	0.2787
8	Mubende	1.10	1.24	0.21810	0.2707	36	Mukono	0.76	0.73	0.21920	0.1926
9	Arua	1.10	1.19	0.18730	0.217	37	Katakwi	0.76	0.69	0.21790	0.2712
10	Sembabule	1.04	1.21	0.26530	0.3861	38	Kanungu	0.75	0.68	0.22140	0.2647
11	Kitgum	1.03	1.15	0.24120	0.3038	39	Soroti	0.74	0.67	0.15300	0.2537
12	Nakapiripirit	1.03	1.20	0.27410	0.3881	40	Kayunga	0.74	0.65	0.20980	0.2585
13	Kisoro	1.02	1.15	0.25920	0.3422	41	Lira	0.74	0.71	0.21520	0.1648
14	Kamwenge	1.00	1.13	0.25650	0.3391	42	Kasese	0.73	0.67	0.17960	0.2091

No.	District	RR-Poisson with gamma prior (DIC=284.461)	RR Poisson with flat prior(DIC =286.274)	sd-gamma prior	sd – flat prior	No.	District	RR-Poisson with gamma prior(DIC= 284.461)	RR Poisson with flat prior(DIC= 286.274)	sd-gamma prior	sd – flat prior
15	Bundibugyo	0.96	1.07	0.26140	0.347	43	Kumi	0.73	0.65	0.20400	0.2471
16	Rakai	0.92	0.97	0.20980	0.2538	44	Rukungiri	0.73	0.63	0.19110	0.2329
17	Nakasongola	0.87	0.91	0.20990	0.3728	45	Luwero	0.72	0.65	0.20170	0.2189
18	Kibaale	0.86	0.88	0.16750	0.2618	46	Ntungamo	0.70	0.62	0.18170	0.2079
19	Kalangala	0.85	0.87	0.26080	0.4228	47	Kabale	0.70	0.62	0.17540	0.1936
20	Apac	0.85	0.87	0.20830	0.1902	48	Moyo	0.69	0.55	0.13040	0.2455
21	Tororo	0.85	0.86	0.21090	0.2523	49	Mbarara	0.69	0.65	0.20720	0.1412
22	Busia	0.84	0.85	0.27410	0.3168	50	Jinja	0.68	0.55	0.18010	0.2202
23	Kyenjojo	0.84	0.86	0.24030	0.263	51	Mpigi	0.68	0.57	0.19000	0.2028
24	Masindi	0.82	0.83	0.20380	0.2402	52	Kabarole	0.67	0.54	0.19220	0.2135
25	Hoima	0.82	0.81	0.21470	0.264	53	Adjumani	0.67	0.50	0.20670	0.234
26	Kiboga	0.81	0.80	0.16320	0.306	54	Yumbe	0.66	0.49	0.19210	0.2144
27	Bushenyi	0.81	0.82	0.23620	0.1847	55	Mayuge	0.64	0.44	0.13090	0.2075
28	Kampala	0.81	0.80	0.19300	0.1954	56	Wakiso	0.54	0.40	0.19290	0.1248

Table 4.8 Continued

CHAPTER FIVE: DISCUSSIONS

In this chapter the study presents discussion of the findings. The chapter provides discussion on the suitability of the models used in the study.

5.1 Discussions of Model Evaluation

Findings show that both the Poisson-gamma, log-normal and the BYM had minimal differences in their performance. Findings indicate that SMR gives unstable results especially when the number of observations (observed count, *y*) gets reduced with increase in number of districts. Subsequent analysis using SAE technique with current increase in the number of districts calls for use of the HB approach which smoothes estimates than the use of SMR.

5.2 Comparison of the Hierarchical Bayes Framework to SMR Results

In 1995 and 2001 there were a total of 38 districts and by 2006 they were 56. Using fewer districts, results show that there were substantial number of deaths (observed count, *y*) of the under-five mortality and hence computations of SMR provided stable estimates. Reliability of SMR was verified using coefficient of variation and showed less variability when few districts are used.

Evaluation of the three model approach (*PG*, Log-Normal and the BYM) revealed feeble difference in performance. It was also observed from the diagnostics test for the presence of autocorrelation that its existence was rather feeble. The absence of a stronger autocorrelation suggests that the Poisson-gamma model could as well perform as good as

the other model options. Additionally, the results presuppose that either the Poisson-gamma or the BYM or the log-normal models appear to be suitable for modeling relative risk of under-five in Uganda.

Findings also show that the Poisson-gamma model potentially perform better than other HB models. The model with the lowest DIC is considered better. It is important to note that the differences between the values of the DIC were generally low indicating that all the three models were fairly good in deriving relative risk of under-five mortality.

5.3 Variation of Relative Risk of Under-five over 1995 to 2006

Results show that the districts of Kalangala, Bundibugyo, Kamuli, Kiboga, Kotido, Mbarara and Kibaale had relatively very high relative risk of under-five mortality by 1995. Kumi district also exhibited a slightly high relative risk of under-five mortality by about 4 percent compared to the national average.

By 2001, results show that the district of Kalangala consistently had a very high SMR and *RR* of under-five mortality. The other districts with very high RR included; Rukungiri, Kisoro, Kabarole. Mpigi, Moyo and Mubende had also a slightly high relative risk of under-five mortality.

Results from the 2006 data show that the districts of Moroto, Kotido, Gulu, Iganga, Kamuli, Pader, Mbale, Mubende and Arua were identified with high relative risk (>=1.10) of under-five mortality. Sembabule, Kitgum, Nakapiripirit, Kisoro and Kamwenge had equally high relative risk of under-five mortality.

Results also show that the districts of Moroto, Kotido, Kamuli, Kisoro, Iganga, Mubende and Kamuli have consistently had high relative risk of under-five mortality over the last one decade. Kalangala district appear to have improved by 2006 compared to the period of 1995 and 2001.

5.4 Spatial Pattern of Relative risk of Under-five Mortality in Uganda and Plausible Causes

This section utilizes literature to explain any linkage with results obtained from the models on the relative risk of under-five mortality.

The districts of Moroto and Kotido as part of Karamoja region found to have the highest relative risk is attributed to high levels of insecurity due to cattle rustling, high poverty levels (66 percent live below poverty line-(UBOS and Macro International, 2007)) and chronic cyclical drought resulting into food-insecurity. As a result, many families are forced to depend on inadequate food supplies from humanitarian agencies like WFP and sometimes from the Office of the Prime Minister (OPM). Inadequate social service delivery in the districts due to inaccessibility, poor quality of services, a mismatch between available services and the traditional dictates of the indigenous pastoral communities together with prevailing and widespread poverty has harshly impacted on the lives of children and the likely cause of high under-five mortality in the districts. The high rates of malnutrition in Karamoja region could partly also be attributed to food insecurity due to drought and frequent floods coupled with the chronic insecurity, and poverty in the region; poor public health system with barriers to health care access; inadequate social and child care environment attributed partly to lack of health and nutrition education leading to poor infant and young child feeding practices (MoH, 2008).

Compared to other districts in the region, the problem in Kotido is exacerbated by the high prevalence of malaria at 48 percent among households which could have been attributed to the low mosquito net coverage of 14 percent (MoH, 2008). The districts are also located in malaria endemic areas.

The districts of Gulu, Pader and Kitgum had problems of insurgency for over 20 years beginning in 1986 and nearly 90 percent of the populations were confined into IDP camps with pathetic conditions of life. Gulu district in particular, under-five mortality rate by 2005 was well above emergency thresholds of 2 per 10 000 per day (MOH, 2005). This was attributed to mainly poor conditions of living, malaria and AIDS was the top self-reported causes of death among children under-five.

Although under-five mortality rate for the South Western part of the country was found to be 181 deaths per 1000 live births much higher than the national average of 137 deaths (UBOS and Macro International, 2007), findings show that it is only the district of Kisoro that had a slightly high relative risk of 1.02. The other districts in the same region of Bushenyi, Kabale, Kanungu, Kisoro, Mbarara, Ntungamo and Rukungiri have enjoyed relatively peace over the years and have fertile soils, reliable rainfall of two planting seasons in a year. The poverty levels are also considerably lower with less than 21 percent compared to the national average of 31 percent by 2006. High *RR* in the district of Kisoro over time is likely to have been attributed to limited accessibility to health services due to the mountainous nature of the area. Kisoro district is located at very high altitude of over 7,000 feet above sea level. Compared to other districts in the country where the highest height above sea level are in the range of about 4000 feet above sea level, Kisoro district remains at a very high altitude. Studies have shown that pneumonia

incidence rates are much higher in high-altitudes than at lower altitudes and likely leading cause of deaths in such areas. There are also many problems associated with malnutrition in the Southwestern region of the country (UBOS and Macro International, 2007). Pneumonia in Kisoro is more likely to be exacerbated by effect of malnutrition.

By 1995, Bundibugyo, Mbarara and Kibale districts had high relative risk of under-five mortality while Rukungiri, Kisoro and Kabarole had the highest by 2001 from the western region. In Bundibugyo district, high under-five mortality is likely to have been associated with the effects of insecurity by 1995 as a result of rebel activities by the Allied Democratic Forces.

Results also show a geographical pattern in the under-five mortality especially so in the Northern part of the country. This geographical pattern requires more resource allocation and effective monitoring of government programs like the Northern Uganda Social Action Fund to reduce under-five mortality in the region.

In Iganga district, pneumonia was one of the major killers of children under 5 years. Findings from Iganga district show that mistreatment with anti-malarials, delays in seeking care and low quality of care for children with fatal pneumonia was killing many children (Källander, et al., 2008). In the same district, there was a community knowledge gap on symptoms and biomedical treatment for pneumonia (Hildenwall, et al., 2007) and findings show that poverty was associated with delay to seek care (Rutebemberwa, 2009). Inappropriate knowledge on causes of pneumonia and signs of non severe pneumonia are likely to interfere with compliance with home care messages (Irimu, Nduati, Wafula, and

Lenja, 2008). Another study in the same region showed that Mbale and Iganga districts had the highest proportion of child deaths among the six selected districts of Uganda. Results showed that diarrhea was the leading (33 percent) cause of deaths in Mbale while malaria (34 percent) and measles (34 percent) were the major causes of under-five mortality in Iganga district (Ntozi and Nakanaabi, 1997). On the other hand, Kamuli district was faced with multi-faceted problems including malnutrition among children, malaria and poor hygiene. Poor hygiene has generally been a big problem in the Eastern part of Uganda and this is supported by other studies like Ssenyonga, Muwonge, Twebaze, and Mutyabule (2009) with high incidences of diarrhea in the region.

In a study carried out in Arua district, several risk factors were identified to be associated with under-five mortality to include; poor accessibility to health services and failures in the health systems in terms of medical staff and facilities to meet child health needs (Akello, Nabiwemba, Zirabamuzaale, and Orach, 2008).

A study conducted in Mubende district established that low utilization of treated mosquito net is the likely cause to high relative risk of the under-five in this area (Mufubenga and Kiwuwa, 2004). However, use of treated mosquito nets is equally becoming a challenge due to reduced susceptibility to some types of malaria vectors. A study by Rubaihayo, Tukesiga, and Abaasa (2008) found out that the pyrethroid insecticide treated nets had a reduced susceptibility to the malaria vector Anopheles gambiae s.l. in Western Uganda.

A study carried out in Sembabule (Mbonye, 2003) revealed that most children with fever, diarrhea, and URTI were treated at home and taken to health units only when they developed life-threatening symptoms. Mbonye further argues that the late referral was complicated by high cost of care, long distances to health units, poor attitude of health workers, lack of drugs at health units, and limited involvement of fathers in care of their children.

Results for the year 1995 and 2001 showed that Kalangala district had the highest *RR* in the country. Findings using UDHS 2006 data show a rather lower *RR* compared to the prior survey results. Kalangala district is an island which had poor accessibility to services like health facilities over time. The district is also located in a high malaria endemic place. By 2006, there appear to have been a big shift arising from a number of factors related to improved infrastructure and level of investments. Before the year 2006, only one ferry was available to the main island of Kalangala through Masaka (Bukakata). However, by the year 2006, another ferry connecting to Kalangala main island was introduced connecting from Entebbe (Kigungu) which is close to the capital city of the country, Kampala. The improved communication is likely to have helped in improved provision of services to Kalangala district. Additionally, BIDCO, a company that heavily invested over US$28 in palm oil is likely to have made a paradigm shift in people's incomes and improved access to services like health. The palm oil investments have also been supported by other donor agencies like the International Fund for Agricultural Development (IFAD). The palm oil investment has resulted in improvement in road infrastructure helping the communities to easily access health facilities especially on the main island. The communities have also been accessing funds through the out growers

scheme to help them in planting palm trees and these funds are likely to have improved people's lives in the area subsequently leading to a likely causal impact in reduced deaths among the children under-five years of age.

Bundibugyo district showed high *RR* by 1995 and by 2006, there was a relatively lower *RR* to under-five mortality. Bundibugyo district is located in the West part of the country and mainly characterized by a mountainous kind of terrain. The problem in Bundibugyo district is likely to have been attributed to difficulty in accessing services and coupled with the Allied Democratic Force rebel insurgency in the 1990s. However, by the year 2003 the rebellion had disappeared and about three-quarters of the total displaced population in the district had returned to their original homes (Hovil and Werker, 2005) and living condition appear to have stabilized including likely causal reduction of the under-fives.

The findings show that results obtained from the HB framework on the spatial pattern of high relative risk of under-five mortality in Uganda show closer associated factors that could be explained.

CHAPTER SIX: SUMMARY OF FINDINGS, CONCLUSION AND RECOMMENDATIONS

In this chapter, the study draws conclusion based on the findings. Recommendations based on estimation of relative risk of under-five mortality using small area estimation techniques are also discussed. The chapter also provides areas of further research that could be under-taken to augment the findings from this work.

6.1 Summary of Findings

The models (Poisson-gamma, log-normal and BYM) show that autocorrelation was not evident and the likely reason why the Poisson-gamma model performed well compared to the log-normal and the BYM the latter two designed to account for autocorrelation.

Sensitivity analysis was conducted to investigate whether results in the models remain stable when a different prior is used i.e robustness of the posterior distribution (Ntzoufras, 2009). Overall results show that after use of the flat and gamma priors there was a minimal difference in DICs. Convergence using the flat prior were reached at a higher value (15,000 iterations) compared to use of a gamma prior (4,000 iterations). Secondly, standard deviations (sd) using a flat prior were slightly higher compared to use of a gamma prior. Overall conclusion derived from the results for *RR* would still remain the same implying robust in the Poisson-gamma model.

Results show that the Poisson-gamma model could provide reliable estimates for relative risk of under-five mortality providing lower DIC. Results also reveal that compared to the modeling approach, utilization of the traditional Standardized Mortality Ratio (SMR) could potentially be associated with high undesirable coefficient of variations (>100%). The modeling approach has added advantage over the commonly used SMR by estimating under-five disease risk for a particular district and smoothening using adjacent district estimates. The study further reveals that it is possible to utilize small area estimation techniques together with national survey data to generate relative risk of under-five mortality for districts in Uganda.

The results further reveal that over the study period 1995 to 2006, the districts of; Kamuli, Kotido, Kisoro and Mubende consistently showed high *RR* of under-five mortality. The district of Kalangala improved by 2006 and this could be attributed to improved communication and more investors (palm growing) with out-growers concept funded by BIDCO and IFAD that started in 2003. The investments meant more resources to the population leading to improved wellbeing hence reducing chances of under-five deaths.

The results also show spatial trend mainly exhibited in the northern region where *RR* is high compared to other regions in Uganda.

6.2 Conclusion

The study demonstrates that use of SMR when used for estimation of relative risk of under-five mortality need to be evaluated using *CV*. Some of the *CVs* were found to be

undesirably high (>100%). This implies that when *CV*s are high, estimation of relative risk of under-five mortality using SMR would be biased.

This study concludes that; with the continued increment on the number of districts in Uganda and when no substantial increment in the sample size is made to the DHS, use of SMR is extremely 'noisy' compared to use of the small area estimation models. The study also show that the hierarchical Bayes approach was suitable for deriving relative risk of under-five mortality in Uganda compared to use of SMR.

This study show that modeling approach through use of; log-normal, Poisson-gamma and the BYM yield better and more stable results than the traditional method of SMR. This study also demonstrates the use of small area estimation techniques to derive relative risk of under-five mortality for districts in Uganda. The study demonstrated that relative risk of under-five mortality can be derived for districts. The study show spatial variations of relative risk of under-5 mortality in Uganda with mainly the northern part of the country having high relative risks. These findings are very important for districts with high RR to justify and mobilize resources to address under-five mortality. Districts with high relative risk of under-five mortality can be targeted for increased resource funding and mobilization.

Furthermore, the study show that with the increase in the number of districts (from 37 in 1995 to 56 by 2006 in the survey data) and yet the sample size has not substantially been increasing, use of SMR becomes extremely unstable compared to the use of HB model approach which borrow strength from the neighboring districts to smoothen the estimate on relative risk of under-five mortality. The study provides a code fragment using

WinBUGS 1.4 software that can be re-used in similar national survey data for Uganda or other countries.

The study showed that autocorrelation in the national surveys for the period 1995, 2001 and 2006 was feeble and the likely reason why results from the Poisson-gamma model was performing as good as other models like the log-normal and the BYM.

The study helped to show the variation over the period 1995, 2001 and 2006 for the relative risk of under-five mortality at district level using UDHS.

The findings from the HB framework and the literature on associated socio-economic factors show closer associations that could explain variability of districts with high relative risk of under-5 mortality in Uganda.

6.3 Recommendations

This study draws key recommendations

1. There is need to evaluate SMR using the coefficient of variation to measure relative risk of under-five mortality
2. Small area estimation need to be used whenever relevant survey data is available to monitor districts with high relative risk of under-five mortality as one way for health interventions in the country.
3. With current increase in the number of districts to more than 110, future national surveys should incorporate at least a sub-county to allow small area estimation at

that level. A proposal to have a sub-county level in survey data is on a premise that a district constitutes at least one sub-county.

4. Current allocation of resources to districts is mainly based on the population and less emphasis on disease burden. Allocation of resources needs to take into account disease burden by taking advantage of small area estimation techniques.
5. This research found out that the Poisson-gamma, log-normal and the BYM were all good in deriving relative risk of under-five mortality in Uganda. Their differences after comparison using DIC were feeble. This implies that the Poisson-gamma which is a basic HB model could potentially be a good estimator of relative risk of under-five mortality in Uganda.

6.4 Areas for Further Research

- Further research could be carried out to explore the hierarchical models incorporating covariates.
- Further research on whether incorporating census data as auxiliary information would yield better results in the estimation of relative risk of under-five mortality. Related to this, further research could explore inclusion of weighted variable in the modeling given the fact that selections of some districts were oversampled and populations are not evenly distributed.
- Spatial differentials may also exist with neonatal deaths as established from other studies (Adebayo, et al., 2004) and again this may be an area of further research in Uganda's case. It may therefore be of interest to establish whether spatial differentials exist differently for neonatal mortality that might call for different intervention strategies. Area of further research using Geoadditive Survival Models appear to offer benefits in taking into account age at deaths of the child

after being applied in a number of countries like Nigeria (Adebayo, et al., 2004; Kandala, et al., 2007) and Malawi (Kandala and Ghilagaber, 2006) on small area estimation for child mortality.

REFERENCES

Adebayo, S. B., & Fahrmeir, L. (2005). Analyzing child mortality in Nigeria with geoadditive discrete-time survival models. *Statistics in Medicine*, 24(5), 709-728.

Adebayo, S. B., Fahrmeir, L., & Klasen, S. (2004). Analyzing infant mortality with geoadditive categorical regression models: a case study for Nigeria. *Econ Hum Biol*, 2(2), 229-244

Akello, B., Nabiwemba, E., Zirabamuzaale, C., & Orach, C. G. (2008). Risk Factors for Perinatal Mortality in Arua Regional Referral Hospital, West Nile, Uganda. *East Afr J Public Heath*, 5, 180-185.

Alexander, N., Moyeed, R., & Stander, J. (2000). Spatial modelling of individual-level parasite counts using the negative binomial distribution. *Biostatistics*, 1(4), 453-463.

Anderson, R. M. (Ed.). (1993). Epidemiology. In Cox, F. E. G. (ed.), *Modern Parasitology*, 2nd Edn. Oxford: Blackwell.

Arul, E., Morgan, G., Mengersen, K., Louise, R., Summerhayes, R., & Beard, J. (2007). Evaluating the effect of neighbourhood weight matrices on smoothing properties of Conditional Autoregressive (CAR) models. *International Journal of Health Geographics*, 6, 54.

Asiimwe, J. B., Jehopio, P., Atuhaire, L. K., & Mbonye, A. K. (2010). Examining small area estimation techniques for public health intervention: Lessons from application to under-five mortality data in Uganda. *Journal of Public Health Policy*, 32(1), 1-15.

Ayiko, R., Antai, D., & Kulane, A. (2009). Trends and determinants of under-five mortality in Uganda. *East Afr J Public Health*, 6(2), 136-140.

Bailey, T. C. (2001). Spatial statistical methods in health. *Cad Saude Publica*, 17(5), 1083-1098.

Berg, A., Meyer, R., & Yu, J. (2004). Deviance Information Criterion for Comparing Stochastic Volatility Models. *Journal of Business and Economic Statistics*, 22(1), 107-120.

Bergamaschi, R., Montomoli, C., Candeloro, E., Monti, M. C., Cioccale, R., Bernardinelli, L., et al. (2006). Bayesian mapping of multiple sclerosis prevalence in the province of Pavia, northern Italy. *Journal of the Neurological Sciences*, 244, 127-131.

Best, N., Cockings, S., Bennett, J., Wakefield, J., & Elliott, P. (2001). Ecological regression analysis of environmental benzene exposure and childhood leukaemia: sensitivity to data inaccuracies, geographical scale and ecological bias. *J R Statist Soc A*, 164, 155-174.

Bolstand, W. M. (2007). *Introduction to Bayesian Statistics* (2nd Edition). Hoboken, Newjersey: John Wiley & Sons, Inc.

Bryce, J., Daelmans, B., Dwivedi, A., Fauveau, V., Lawn, J. E., Mason, E., et al. (2008). Countdown to 2015 for maternal, newborn and child survival: the 2008 report on tracking coverage of interventions. *Lancet*, 371, 1247-1258.

Catelan, D., Lagazio, C., & Biggeri, A. (2010). A hierarchical Bayesian approach to multiple testing in disease mapping. *Biometrical Journal*, 52(6), 784-797.

Chen, G. (2009). *Essays on Model Selection Using Bayesian Inference*. Unpublished PhD, Rutgers, The State University of New Jersey, New Brunswick, New Jersey.

Chowdhury, H. Q., Islam, R., & Hossain, K. (2009). Levels and Patterns Of Child Mortality: Application Of Indirect Techniques. *The Internet Journal of Health*, 9(1).

Clayton, D., & Bernardinelli, L. (1996). *Geographic and Environmental Epidemiology: Methods for small area studies*. In C. J. Elliot P., English D., Stern R. (Ed.), Bayesian methods for mapping disease risk (pp. 205-220). UK Oxford University Press.

Clayton, D., & Kaldor, J. (1987). Empirical Bayes Estimates of Age-Stardardized relative risks for use in disease mapping. *Biometrics*, 43, 671-681.

Cressie, N. (1992). Smoothing regional maps using empirical Bayes predictors. *Geograph. Anal.* , 24, 75-95.

Cressie, N., & Chan, N. H. (1989). Spatial modeling of regional variables. *J. Amer. Statist. Assoc.*, 84, 393-401.

Downing, A., Forman, D., Gilthorpe, M. S., Edwards, K. L., & Manda, S. O. M. (2008). Joint Disease Mapping using Six Cancers in the Yorkshire Region of England. *International Journal of Health Geographics*, 7, 41.

Earnest, A., Beard, J. R., Morgan, G., Lincoln, D., Summerhayes, R., Donoghue, D., et al. (2010). Small area estimation of sparse disease counts using shared component models-application to birth defect registry data in New South Wales, Australia. *Health & Place*, 16(4), 684-693.

Earnest, A., Morgan, G., Mengersen, K., Ryan, L., Summerhayes, R., & Beard, J. (2007). Evaluating the effect of neighbourhood weight matrices on smoothing properties of Conditional Autoregressive (CAR) models. *International Journal of Health Geographics*, 6, 54.

Freund, J. E. (1992). *Mathematical Statistics* (5th Edition). Newjersey: Prentice-Hall, Inc.

Gangnon, R. E., & Clayton, M. K. (2003). A hierarchical model of spatially clustered disease rates. *Statistics in Medicine*, 22, 3213-3228.

Goodman, M. S. (2010). Comparison of small-area analysis techniques for estimating prevalence by race. *Preventive Chronic Disease*, 7(2).

Goovaerts, P., & Gebreab, S. (2008). How does Poisson kriging compare to the popular BYM model for mapping disease risks? *International Journal of Health Geographics*, 7, 6.

Graham, P. (2008). Intelligent Smoothing Using Hierarchical Bayesian Models. *Epidemiology*, 19(3), 493-495.

Haininga, R., Lib, G., Maheswaranc, R., Blangiardod, M., Lawe, J., Bestd, N., et al. (2010). Inference from ecological models: Estimating the relative risk of stroke from air pollution exposure using small area data. *Spatial and Spatio-temporal Epidemiology*, 1(2), 123-131.

Hildenwall, H., Rutebemberwa, E. K., Nsabagasani, X., Pariyo, G., Tomson, G., & Peterson, S. (2007). Local illness concepts-Implications for management of childhood pneumonia in Eastern Uganda. *Acta Trop*, 101, 217-224.

Hill, K., & Figueroa, M. (2001). Essays in Medical Demography. In B. Zaba & J. Blacker (Eds.), *Child Mortality Estimation by Time since First Birth*. London & New York: The Athlone Press.

Holsinger K. (2012). The Deviance Information Criterion. Retrieved 3rd October 2010, from http://darwin.eeb.uconn.edu/eeb348/lecture-notes/testing-hardy-einberg/node5.html.

Hovil, L., & Werker, E. (2005). Portrait of a Failed Rebellion: An Account of Rational, Sub-optimal Violence in Western Uganda. *Rationality and Society*, 17, 5-34.

Irimu, G., Nduati, R. W., Wafula, E., & Lenja, J. (2008). Community understanding of pneumonia in Kenya. *African Health Sciences*, 8(2), 103-107.

Jia, H., Muennig, P., & Borawski, E. (2004). Comparison of small-area analysis techniques for estimating county-level outcomes. *American Journal of Preventive Medicine*, 26(5), 453-460.

Johnson, G. D. (2004). Small Area Mapping of Prostate Cancer Incidence in New York State (USA) using fully Bayesian Hierarchical Modelling. *International Journal of Health Geographics*, 3, 29.

Julious, S. A., Nicholl, J., & George, S. (2001a). Why do we continue to use standardized mortality ratios for small area comparisons? *J Public Health Med*., 23, 40-46.

Julious, S. A., Nicholl, J., & George, S. (2001b). Why do we continue to use standardized mortality ratios for small area comparisons? *Journal of Public Health Medicine*, 23(1), 40-46.

Kahn, J. M., Kramer, A. A., & Rubenfeld, G. D. (2007). Transferring Critically Ill Patients Out of Hospital Improves the Standardized Mortality Ratio: A Simulation Study. *Chest*, 131(1), 68-75.

Källander, K., Hildenwall, H., Waiswa, P., Galiwango, E., Peterson, S., & Pariyo, G. (2008). Delayed care seeking for fatal pneumonia in children aged under five years in Uganda: a case-series study. *Bulletin of the World Health Organization*, 86(5), 333.

Kandala, N., & Ghilagaber, G. (2006). Geo-Additive Bayesian Discrete-Time Survival Model and its Application to Spatial Analysis of Childhood Mortality in Malawi. *Quality & Quantity*, 40(6), 935-957.

Kandala, N., Ji, C., Stallard, N., Stranges, S., & Cappuccio, F. P. (2007). Spatial Analysis of Risk Factors for Childhood Morbidity in Nigeria. *American Journal of Tropical Medicine and Hygiene*, 77(4), 770-778.

Khatab, K., & Fahrmeir, L. (2009). Analysis of childhood morbidity with geoadditive probit and latent variable model: a case study for Egypt. *The American Society of Tropical Medical Hygiene*, 81(1), 116-128.

Kleinschmidta, I., Bagayokob, M., Clarkec, G. P. Y., Craiga, M., & Le-Sueura, D. (2000). A spatial statistical approach to malaria mapping. *International Journal of Epidemiology*, 29(2), 355-361.

Knutson, K., Zhang, W., & Tabnak, F. (2008). Applying the small-area estimation method to estimate a population eligible for breast cancer detection services. Retrieved 3rd October 2010, from http://www.cdc.gov/pcd/issues/2008/jan/06_0144.htm

Korenromp, E. L., Arnold, F., Williams, B. G., Nahlen, B. L., & Snow, R. W. (2004). Monitoring trends in under-5 mortality rates through national birth history surveys. *International Journal of Epidemiology*, 33, 1293-1301.

Lambert, D. (1992). Zero inflated Poisson regression, with an application to defect in manufacturing. *Technometrics*, 34, 1-14.

Last, J. M. (2001). *A Dictionary of Epidemiology*, 4th Edition. New York: Oxford University Press.

Law, D. C. G., Serre, M. L., Christakos, G., & Leone, P. A. (2004). Spatial analysis and mapping of sexually transmitted diseases to optimise intervention and prevention strategies. *Sexually Transmitted Infections*, 80(4), 294-299.

Lawson, A. B., Biggeri, A. B., Boehning, D., Lesaffre, E., Viel, J. F., Clark, A., et al. (2000). Disease Mapping Models: An Empirical Evaluation. *Statistics in Medicine*, 19, 2217-2241.

Lawson, A. B., Browne, W. J., & Rodeiro, C. L. V. (2003). *Disease Mapping with WinBUGS and MLwiN*. England: John Wiley and Sons, Ltd.

Lawson, A. B., & Williams, F. L. R. (2001). *An introduction Guide to Disease Mapping*. West Sussex, UK: John Wiley & Sons, Ltd.

Lawson, A. W. (1994). Using Spatial Gaussian priors to model heterogeneity in environmental epidemiology. *Royal Statistical Society*, 43, 69-76.

Leyland, A. H., & Davies, C. A. (2005). Empirical Bayes methods for disease mapping. *Statistical Methods in Medical Research*, 14(1), 17-34.

López-Abente, G., Aragonés, N., Pérez-Gómez, B., Ramis, R., Vidal, E., García-Pérez, J., et al. (2008). Kidney cancer mortality in Spain: geographic patterns and possible hypotheses. *BMC Cancer*, 8, 293.

Lord, D. (2006). Modeling motor vehicle crashes using Poisson-gamma models: Examining the effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter. *Accident Analysis and Prevention*, 38(4), 751-766.

Lutambi, A. M., Alexander, M., Jensen, C., Mahutanga, C., & Nathan, R. (2010). Spatial-temporal clusters in Ifakara HDSS in South-eastern Tanzania. Retrieved 21st September 2010, from http://www.globalhealthaction.net

MacNab, Y. C., Farrell, P. J., Gustafson, P., & Wen, S. (2004). Estimation in Bayesian Mapping. *Biometrics*, 60, 865-873.

Maheswaran, R., Haining, R., Pearson, T., Law, J., Brindley, P., & Best, N. (2006). Outdoor and stroke mortality: adjusting for small area level smoking prevalence using a Bayesian approach. *Statistical Methods in Medical Research*, 15(5), 499-516.

Maiti, T. (1998). Hierarchical Bayes Estimation of mortality rates for disease mapping. *Journal of Statistical Planning and Inference*, 69, 339-348.

Manton, K. G., Woodbury, M. A., Stallard, E., Riggan, W. B., Creason, J. P., & Pellom, A. C. (1989). EmpiricalBayes procedures for stabilizing maps of US cancer mortality rates. *J. Amer. Statist. Assoc*., 84, 637-650.

Mbonye, A. K. (2003). Prevelance of Childhood Illness and Care-Seeking Practices in Rural Uganda. *The Scientific World Journal*, 3, 721-730.

Meza, J. L. (2002). Empirical Bayes estimation smoothing of relative risks in disease mapping. *Journal of Statistical Planning and Inference*, 112, 43-62.

Meza, J. L. (2003). Empirical Bayes estimation smoothing of relative risks in disease mapping. *Journal of Statistical Planning and Inference*, 112, 43-62.

MOH. (2005). *Health and mortality survey among internally displaced persons in Gulu, Kitgum and Pader districts, Northern Uganda*. Kampala, Uganda: Ministry of Health (MoH).

MoH. (2008). *Nutrition and Health Assessment in Karamoja Region*. Kampala, Uganda: Ministry of Health (MoH).

Mok, C. C., Kwok, C. L., Ho, L. Y., Chan, P. T., & Yip, S. F. (2011). Life expectancy, standardized mortality ratios, and causes of death in six rheumatic diseases in Hong Kong, China. *Arthritis & Rheumatism*, 63(5), 1182–1189.

Mollie, A. (1995). Markov Chain Monte Carlo in Practice. In W. R. Gilks, S. Richardson & D. J. Spiegelhalter (Eds.), *Bayesian mapping of disease* (pp. 359-379): Chapman & Hall.

Moore, L., Hanley, J. A., Turgeon, A. F., & Lavoie, A. (2012). Comparing regression-adjusted mortality to standardized mortality ratios for trauma center profiling. J *Emerg Trauma Shock*., 5(4), 333-337.

Mufubenga, P., & Kiwuwa, M. S. (2004). *Utilizing community traditional initiatives in Prevention of malaria pregnancy: A path finder for better Antenatal care utilization IPT uptake and ITN use-A randomized community trial baseline survey*. Kampala, Uganda: MoH/WHO.

Ntozi, J. P. M., & Nakanaabi, I. (1997). The AIDS Epidemic and Infant and Child Mortality in Six Districts of Uganda. *Health Trans Rev*, 7, 189-205.

Ntzoufras, I. (2009). *Bayesian Modeling using WinBUGS*. New Jersey: John Wiley & Sons, Inc.

Nuwaha, F., & Mukulu, A. (2009). Trends inunder-fivemortalityinUganda 1954-2000: can Millennium Development Goals be met? *African Health Sciences*, 9(2), 125-128.

Ocaña-Riola, R. (2007). The misuse of count data aggregated over time for disease mapping. *Statistics In Medicine*, 26(24), 4489-4504.

Rajaratnam, J. K., N., T. L., Lopez, A. D., & Murray, C. J. L. (2010). Measuring Under-five Mortality: Validation of New Low-Cost Methods. *PLoS Medicine*, 7(4), 1-24.

Ramis-Prieto, R., Garcı´a-Pe´rez, J., Polla´n, M., Aragone´s, N., Pe´rez-Go´mez, B., & Lo´pez-Abente, G. (2007). Modeling of Municipal mortality due to hematological neoplasias in Spain. *J Epidemiol Community Health*, 61, 165-171.

Randramanana, R. V., Richard, V., Rakotomanana, F., Sabatier, P., & Bicont, D. J. (2010). Bayesian Mapping of Pulmonary Tuberclosis in Antananarivo. Madascar. *BMC Infectious Diseases*, 10, 21.

Rao, J. N. K. (2003). *Small Area Estimation*. New Jersey: John Wiley & Sons, Inc.

Rodeiro, C. L., & Lawson, A. B. (2004). An evaluation of the edge effects in disease map modeling. *Computational Statistics & Data Analysis*, 49, 45 - 62.

Rubaihayo, J., Tukesiga, E., & Abaasa, A. (2008). Reduced susceptibility to pyrethroid insecticide treated nets by the malaria vector Anopheles gambiae s.l. in Western Uganda. *Malaria Journal*, 7, 92.

Rudan, I., Chan, K. Y., Zhang, J. S., Theodoratou, E., Feng, X. L., Salomon, J. A., et al. (2010). Causes of deaths in children younger than 5 years in China in 2008. *Lancet*, 375(9720), 1083-1089.

Rutebemberwa, E. K. (2009). *Access to Health Care for Febrile Children in Uganda*. Unpublished PhD. Makerere University and Karolinska Institutet, Kampala and Sweden.

SAS. (2010). Bayesian Hierarchical Poisson Regression Model for Overdispersed Count Data SAS Institute Inc. Retrieved 11th October 2010, from http://support.sas.com/rnd/app/examples/stat/BayesSalm/new_example/index.html

Schneider, K. L., Lapane, K. L., Clark, M. A., & Rakowski, W. (2009). Using Small-Area Estimation to Describe County-Level Disparities in Mammography. *Preventing Chronic Disease*, 6(4).

Soe, M. M., & Sullivan, K. M. (2006). Standardized Mortality Ratio and Confidence Interval. Retrieved 5th November 2012, from http://www.openepi.com/Documentation/SMRdoc.htm

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Der-Linde, A. V. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), 583-639.

Spiegelhalter, D. J., Thomas, A., Best, N., & Lunn, D. (2003). WinBUGS User Manual Version 1.4. Retrieved 4th February 2009, from http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/manual14.pdf

Ssenyonga, R., Muwonge, R., Twebaze, F. B. N., & Mutyabule, R. (2009). Determinants of acute diarrhoea in children aged 0 - 5 years in Uganda. *East African Medical Journal*, 86(11).

Symons, M. J., & Taulbee, J. D. (1980). Standardized Mortality Ratio as an Approximation to Relative Risk

Retrieved 10th September, 2011, from http://www.stat.ncsu.edu/information/library/mimeo.archive/ISMS_1980_1294.pdf

UBOS, & Macro International. (2007). *Uganda Demographic and Health Survey 2006*: Calverton, Maryland, USA: Uganda Bureau of Statistics (UBOS) and Macro International.

UN, (1983). *Indirect Techniques for Demographic Estimation*, Manual X. New York, United Nations (UN).

UNICEF. (2007). *The State of World's Children 2008*. USA: United Nations Children's Fund (UNICEF).

Wakefield, J. (2007). Disease Mapping and Spatial Regression with Count Data. *Biostatistics*, 8(2), 158-183.

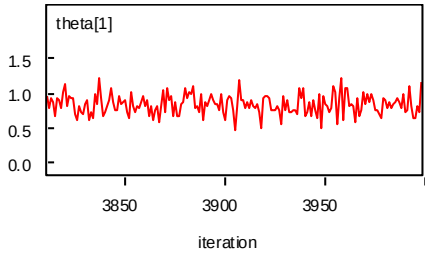
Waldhoer, T., Wald, M., & Heinzl, H. (2008). Analysis of the spatial distribution of infant mortality by cause of death in Austria in 1984 to 2006. *International Journal of Health Geographics*, 7, 21.

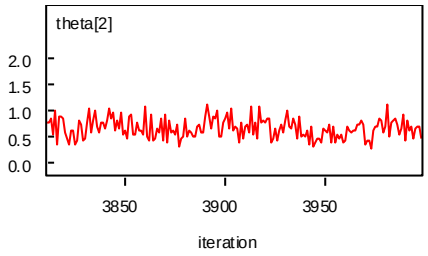
Walsh, L., Dufey, F., Tschense, A., Schnelzer, M., Sogl, M., & Kreuzer, M. (2012). Prostate cancer mortality risk in relation to working underground in the Wismut cohort study of German uranium miners, 1970–2003. BMJ Open Retrieved 12th December, 2012, from http://bmjopen.bmj.com/content/2/3/e001002.full?ga=w_bmjj_bmj-com

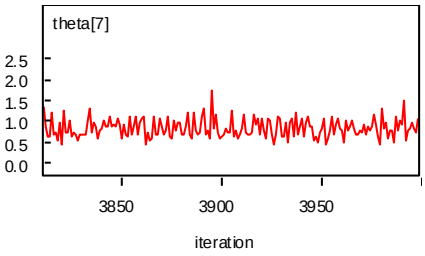
WB. (2008). *Analysis Health Equity using Household Survey Data. A Guide to Technique and their Implementation*. Washington, DC: The World Bank (WB).

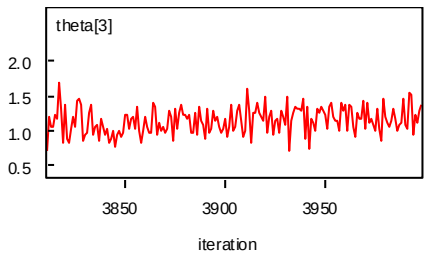
Zhu, L., Carlin, B. P., English, P., & Scalf, R. (2000). Hierarchical Modeling of Spatio-Temporally Misaligned Data: Relating Traffic Density to Pediatric Asthma Hospitalizations. *Environmetrics*, 11, 43-61.

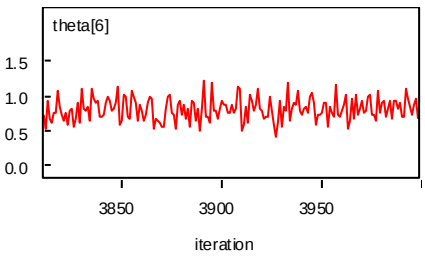
APPENDIX 1: Dynamic Trace Plots for Lognormal Model Using UDHS2006 Data

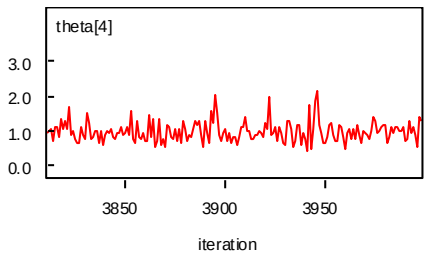


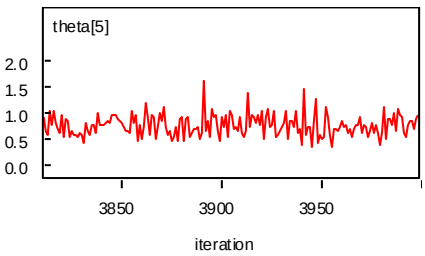


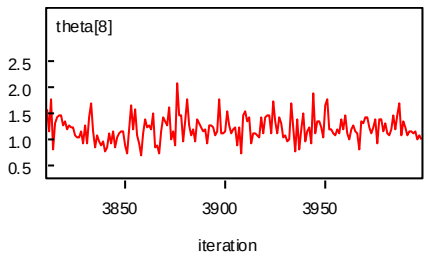


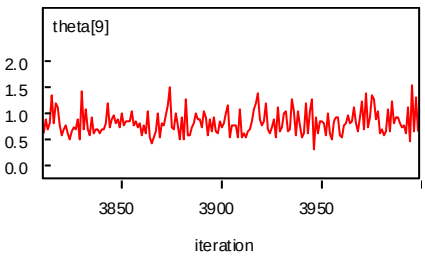


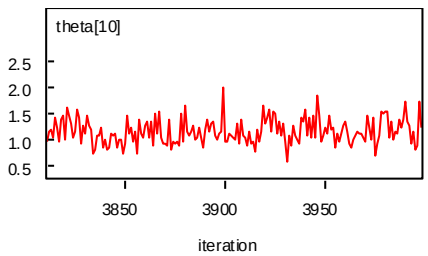


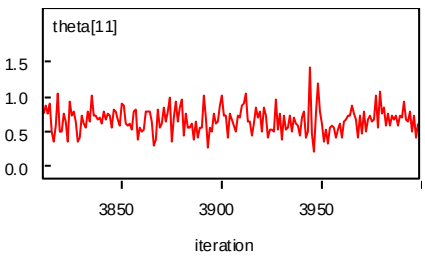


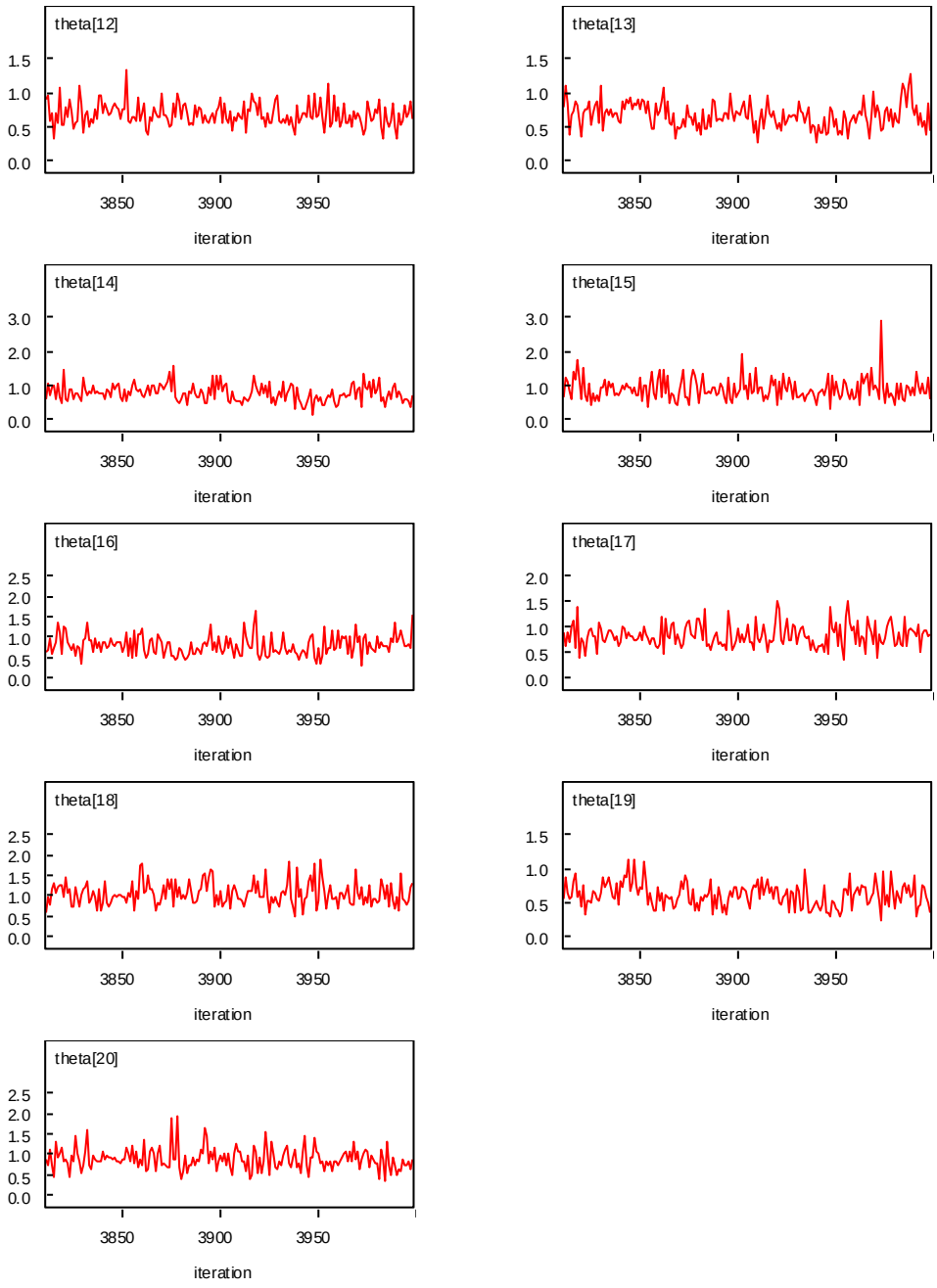




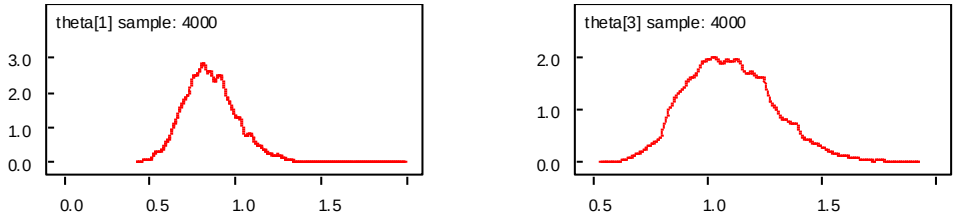


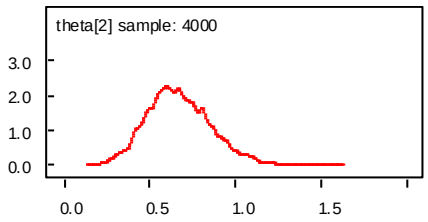


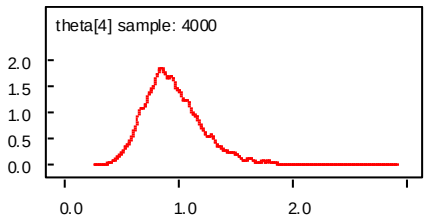


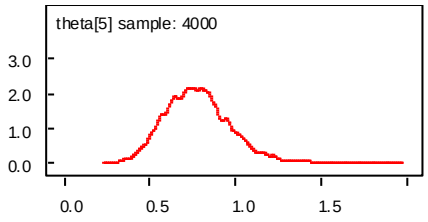


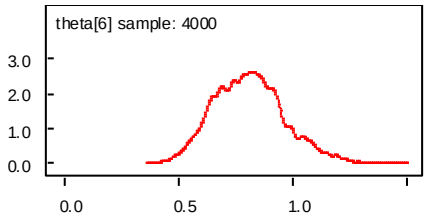
APPENDIX 2: Kernel Density Plots for Lognormal Model Using UDHS2006 Data

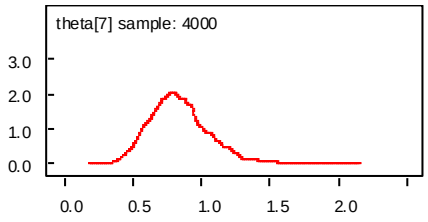


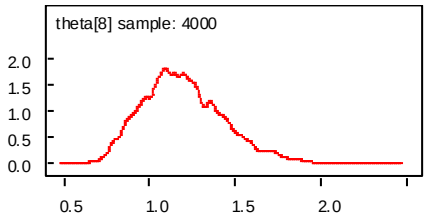

theta[2] sample: 4000
3.0
2.0
1.0
0.0
0.0
0.5
1.0
1.5

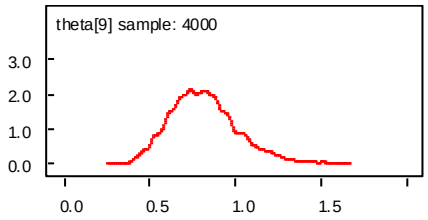

theta[4] sample: 4000
2.0
1.5
1.0
0.5
0.0
0.0
1.0
2.0

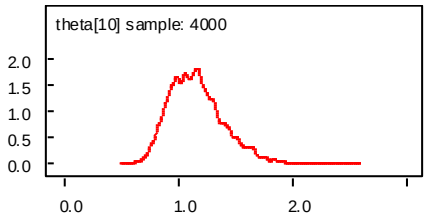

theta[5] sample: 4000
3.0
2.0
1.0
0.0
0.0
0.5
1.0
1.5

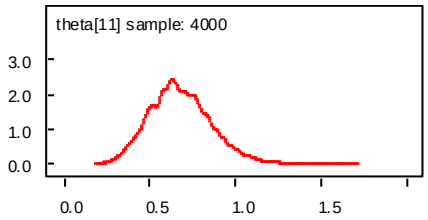

theta[6] sample: 4000
3.0
2.0
1.0
0.0
0.0
0.5
1.0

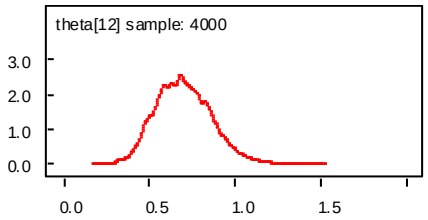

theta[7] sample: 4000
3.0
2.0
1.0
0.0
0.0
0.5
1.0
1.5
2.0

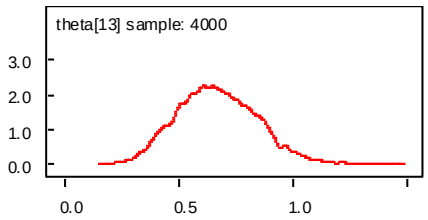

theta[8] sample: 4000
2.0
1.5
1.0
0.5
0.0
0.5
1.0
1.5
2.0

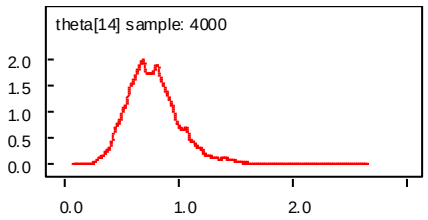

theta[9] sample: 4000
3.0
2.0
1.0
0.0
0.0
0.5
1.0
1.5

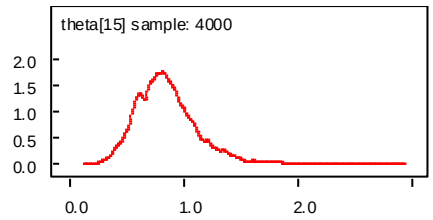

theta[10] sample: 4000
2.0
1.5
1.0
0.5
0.0
0.0
1.0
2.0

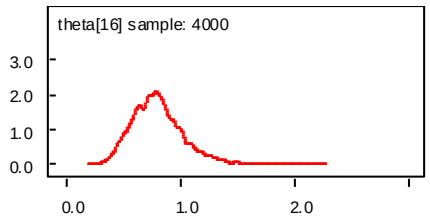

theta[11] sample: 4000
3.0
2.0
1.0
0.0
0.0
0.5
1.0
1.5

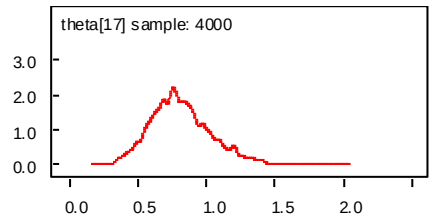

theta[12] sample: 4000
3.0
2.0
1.0
0.0
0.0
0.5
1.0
1.5

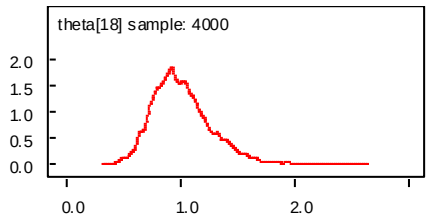

theta[13] sample: 4000
3.0
2.0
1.0
0.0
0.0
0.5
1.0

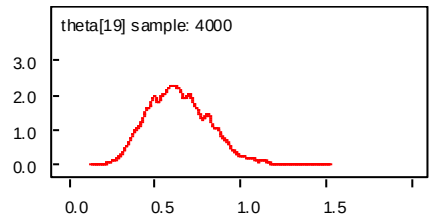

theta[14] sample: 4000
2.0
1.5
1.0
0.5
0.0
0.0
1.0
2.0

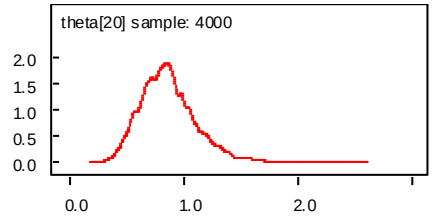




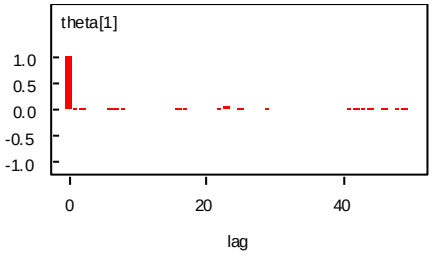


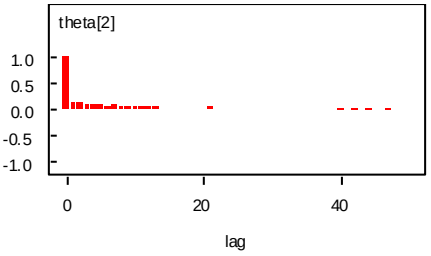


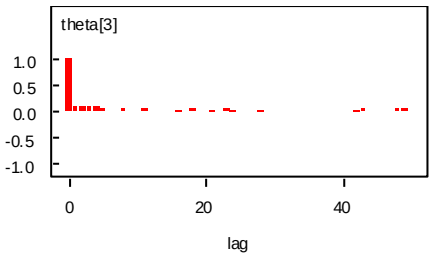


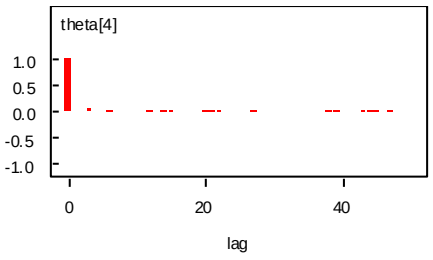


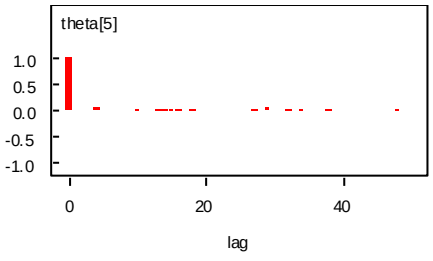
APPENDIX 3: Autocorrelation Plots for Lognormal Model Using UDHS2006 Data

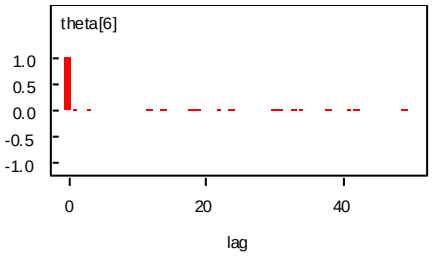


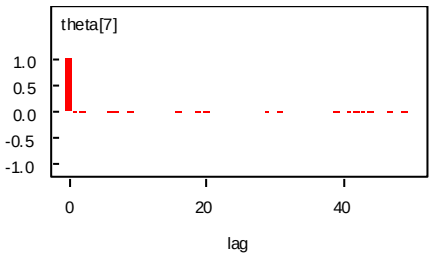


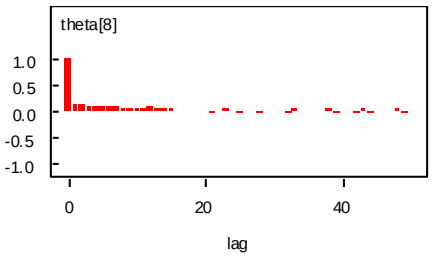

theta[3]
1.0
0.5
0.0
-0.5
-1.0
0
20
40
lag

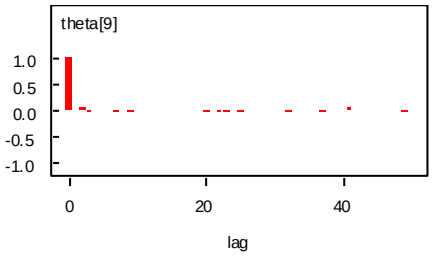

theta[4]
1.0
0.5
0.0
-0.5
-1.0
0
20
40
lag

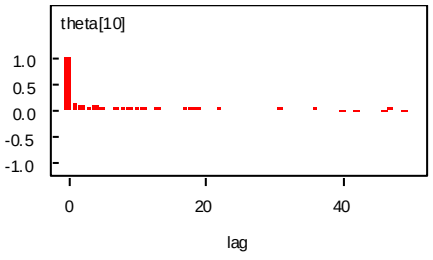

theta[5]
1.0
0.5
0.0
-0.5
-1.0
0
20
40
lag

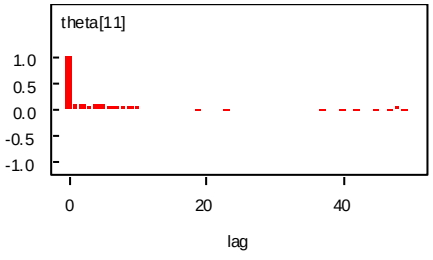

theta[6]
1.0
0.5
0.0
-0.5
-1.0
0
20
40
lag

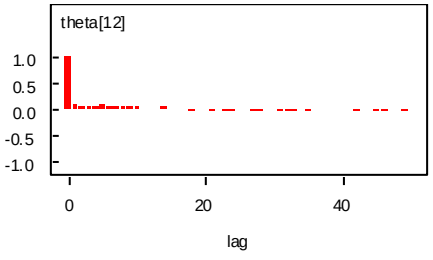

theta[7]
1.0
0.5
0.0
-0.5
-1.0
0
20
40
lag

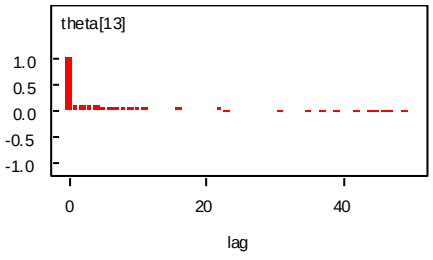

theta[8]
1.0
0.5
0.0
-0.5
-1.0
0
20
40
lag

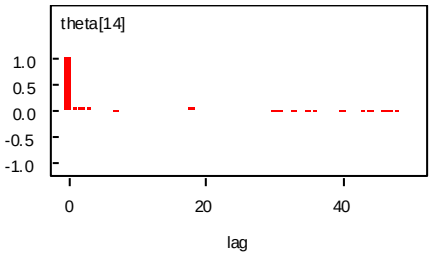

theta[9]
1.0
0.5
0.0
-0.5
-1.0
0
20
40
lag

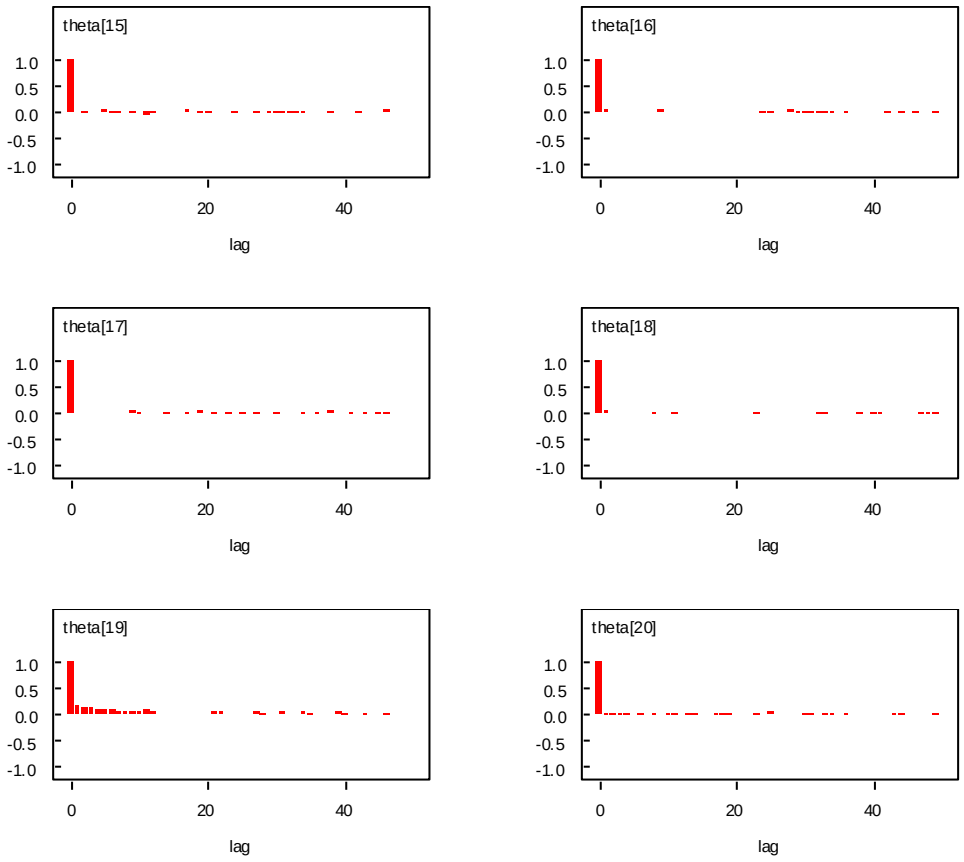

theta[10]
1.0
0.5
0.0
-0.5
-1.0
0
20
40
lag


theta[11]
1.0
0.5
0.0
-0.5
-1.0
0
20
40
lag


theta[12]
1.0
0.5
0.0
-0.5
-1.0
0
20
40
lag


theta[13]
1.0
0.5
0.0
-0.5
-1.0
0
20
40
lag


theta[14]
1.0
0.5
0.0
-0.5
-1.0
0
20
40
lag



APPENDIX 4: Code Used in WinBUGS1.4 Software

********* Poisson-gamma model

model
{
for (i in 1:N){
y[i] ~ dpois(mu[i])
mu[i]<-e[i]*theta[i]
theta[i]~dgamma(a,b)
}

a~dexp(0.1)
b~dexp(0.1)

mean<-a/b
var<-a/pow(b,2)
}

********* log-normal model

```
model
        {
                        for (i in 1 : N) {
                                        y[i] ~ dpois(mu[i])
                                        log(mu[i]) <- log(e[i])+alpha+v[i]
                                        theta[i]<-exp(alpha+v[i])
                                        v[i]~dnorm(0,tau)
                        }
                        tau~dgamma(0.5,0.0005)
                        alpha~dnorm(0,1.0E-5)
                        mean<-exp(alpha)
 }
```

***Figure deflated by a constant factor for comparability

```
list(e=c(20.061,5.251,21.646,4.465,7.741,19.413,4.224,13.939,7.307,13.303,7.275,12.165,7.589,2.240,
0.811,3.239,4.341,5.848,6.094,2.404,16.392,13.277,5.608,5.434,11.122,5.085,5.573,8.631,8.282,
15.103,6.629,8.020,21.740,9.054,17.969,9.771,12.218,28.888,4.847,5.056,9.508,13.044,15.047,3.942,
11.300,10.086,9.571,8.856,10.967,7.304,4.198,4.816,6.291,9.133,21.171,6.535),
y=c(17.4,1.0,27.3,5.7,5.2,15.6,3.5,21.0,5.7,19.2,2.6,6.5,2.6,0.9,0.7,1.9,3.2,8,0.8,2.3
,12.7,19.2,7.4,2.9,6.6,2.7,2.7,7.7,10.8,24.1,3.4,6.8,14.7,4.9,14.0,7.9,16.7,17.7,11.1,1.3,4.1,17.7,10.4,6.1,8.7,5.2,13.9,6.5
,11.1,3.6,6.6,2.6,3.4,7.8,6.3,1.6),N=56)
```

********* Using the BYM model

```
model
        {
                        for (i in 1 : N) {
                                        y[i] ~ dpois(mu[i])
                                        log(mu[i]) <- log(e[i])+alpha+u[i]+v[i]
                                        theta[i]<-exp(alpha+u[i]+v[i])
                                        v[i]~dnorm(0,tau.v)
                        }
                        eps<-1.0E-6
                        u[1:56]~car.normal(adj[],weights[],num[],tau.u)
        for(k in 1:SN)
                        {
                                weights[k]<-1
                        }
                        alpha~dflat()
                        mean<-exp(alpha)
                        tau.u~dgamma(0.5,0.0005)
                        tau.v~dgamma(0.5,0.0005)
 }
```

```
list(N=56, SN=276,
e=c(4.847,15.103,13.939,13.303,13.277,9.571,12.218,13.044,21.646,4.198,
```

8.282,3.942,5.848,5.608,4.465,10.967,2.404,8.631,0.811,20.061,
9.133,4.224,8.02,9.771,7.307,4.341,19.413,16.392,11.3,17.969,
3.239,8.856,2.24,7.741,4.816,15.047,5.085,5.434,6.291,5.573,
21.74,11.122,6.629,7.304,9.054,10.086,12.165,5.056,28.888,7.275,
9.508,7.589,5.251,6.535,21.171,6.094),
y=c(11.1,24.1,21,19.2,19.2,13.9,16.7,17.7,27.3,6.6,
10.8,6.1,8,7.4,5.7,11.1,2.3,7.7,0.7,17.4,
7.8,3.5,6.8,7.9,5.7,3.2,15.6,2.7,8.7,14,
1.9,6.5,0.9,5.2,2.6,10.4,2.7,2.9,3.4,2.7,
14.7,6.6,3.4,3.6,4.9,5.2,6.5,1.3,17.7,2.6,
4.1,2.6,1,1.6,6.3,0.8),
adj=c(2,41,37,12,
11,6,41,1,
53,9,29,24,20,6,11,
5,50,55,34,21,32,
41,40,50,4,32,39,33,
11,3,20,41,2,
35,43,32,21,31,
26,18,23,49,10,51,56,45,
54,29,3,53,48,
8,23,49,16,30,51,
3,6,2,
1,37,43,35,31,
38,47,
23,52,42,27,49,
25,42,52,23,18,
10,49,30,19,
20,24,45,40,41,
25,15,52,23,8,26,
56,51,30,16,36,
6,3,24,17,41,
32,4,34,22,7,
21,34,
18,15,52,14,49,10,8,
3,29,25,26,45,17,20,
24,15,18,26,
24,25,18,8,45,
14,42,44,46,49,
56,36,
9,24,3,
51,10,16,19,
12,35,7,
43,39,5,4,21,7,35,
41,5,39,37,
21,4,55,22,
12,43,32,7,31,
50,40,45,56,28,19,55,
1,41,33,39,43,12,
44,13,47,
37,33,5,32,43,
41,17,45,36,50,5,
6,20,17,40,5,33,37,1,2,
52,15,27,14,
37,39,32,7,35,12,
27,38,47,46,
17,24,26,8,56,36,40,
27,44,47,49,
46,44,38,13,
54,9,53,
23,14,27,46,16,10,8,
4,5,40,36,55,
8,10,30,19,56,
15,42,14,23,18,
48,54,9,3,
9,53,48,
4,50,36,34,
28,45,8,51,19,36),
num=c(4,4,7,6,7,5,5,8,5,6,
3,5,2,5,5,4,5,6,5,5,
5,2,7,7,4,5,5,2,3,4,
3,7,4,4,5,7,6,3,5,6,
9,4,6,4,7,4,4,3,7,5,

```
5,5,4,3,4,6))
list(alpha=0,tau.v=1, tau.u=1,
u=c(0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0),
v=c(0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0))
```

**********Approach 2: Poisson-gamma using diffuse prior

```
model
{
for (i in 1:N){
                    y[i] ~ dpois(mu[i])
                    mu[i]<-e[i]*theta[i]
                    theta[i]~dgamma(a,b)
                    }
a~dgamma(1,1)
b~dgamma(1,1)

}
```

APPENDIX 5: Computation of Expected Value

No.	District	Unweighted under-five deaths(y_i) - derived from DHS2006	Population 2002	2006 Under-five mortality per 1000 live birth	Region	Population Projection (2006)	Estimated Population Under five	**Expected death under-five**
1	Apac	174	683,993	200	North/IDP Area	783,638	150,459	**30,092**
2	Adjumani	10	202,290	177	West Nile/North West	231,760	44,498	**7,876**
3	Arua	273	833,928	177	West Nile/North West	955,416	183,440	**32,469**
4	Budibugyo	57	209,978	145	Western	240,568	46,189	**6,697**

No.	District	Unweighted under-five deaths(y_i) - derived from DHS2006	Population 2002	2006 Under-five mortality per 1000 live birth	Region	Population Projection (2006)	Estimated Population Under five	**Expected death under-five**
5	Bugiri	52	412,395	128	East Central	472,473	90,715	**11,612**
6	Bushenyi	156	731,392	181	South Western	837,943	160,885	**29,120**
7	Busia	35	225,008	128	East Central	257,788	49,495	**6,335**
8	Gulu	210	475,260	200	North/IDP Area	544,497	104,543	**20,909**
9	Hoima	57	343,618	145	Western	393,677	75,586	**10,960**
10	Iganga	192	708,690	128	East Central	811,933	155,891	**19,954**
11	Jinja	26	387,573	128	East Central	444,035	85,255	**10,913**
12	Kabale	65	458,318	181	South Western	525,087	100,817	**18,248**
13	Kabarole	26	356,914	145	Western	408,910	78,511	**11,384**
14	Kaberamaido	9	131,650	116	Eastern	150,829	28,959	**3,359**
15	Kalangala	7	34,766	159	Central 1	39,831	7,648	**1,216**
16	Kapchorwa	19	190,391	116	Eastern	218,128	41,880	**4,858**
17	Kiboga	32	229,472	129	Central 2	262,902	50,477	**6,512**
18	Kisoro	8	220,312	181	South Western	252,407	48,462	**8,772**
19	Mayuge	8	324,674	128	East Central	371,973	71,419	**9,142**
20	Nakasongola	23	127,064	129	Central 2	145,575	27,950	**3,606**
21	Kampala	127	1,189,142	94	Kampala	1,362,378	261,577	**24,588**
22	Kamuli	192	707,332	128	East Central	810,377	155,592	**19,916**
23	Kamwenge	74	263,730	145	Western	302,151	58,013	**8,412**
24	Kanungu	29	204,732	181	South Western	234,558	45,035	**8,151**
25	Kasese	66	523,033	145	Western	599,229	115,052	**16,683**
26	Katakwi	27	298,950	116	Eastern	342,502	65,760	**7,628**
27	Kayunga	27	294,613	129	Central 2	337,533	64,806	**8,360**
28	Kibaale	77	405,882	145	Western	465,012	89,282	**12,946**
29	Kitgum	108	282,375	200	North/IDP Area	323,512	62,114	**12,423**
30	Kotido	241	591,889	174	North/ Karamoja Region	678,117	130,198	**22,655**
31	Kumi	34	389,665	116	Eastern	446,432	85,715	**9,943**
32	Kyenjojo	68	377,171	145	Western	432,118	82,967	**12,030**
33	Lira	147	741,240	200	North/IDP Area	849,225	163,051	**32,610**
34	Luwero	49	478,595	129	Central 2	548,318	105,277	**13,581**
35	Masaka	140	770,662	159	Central 1	882,933	169,523	**26,954**

No.	District	Unweighted under-five deaths(y_i) - derived from DHS2006	Population 2002	2006 Under-five mortality per 1000 live birth	Region	Population Projection (2006)	Estimated Population Under five	**Expected death under-five**
36	Masindi	79	459,490	145	Western	526,429	101,074	**14,656**
37	Mbale	167	718,240	116	Eastern	822,875	157,992	**18,327**
38	Mbarara	177	1,088,356	181	South Western	1,246,910	239,407	**43,333**
39	Moroto	111	189,940	174	North/ Karamoja Region	217,611	41,781	**7,270**
40	Moyo	13	194,778	177	West Nile/North West	223,154	42,845	**7,584**
41	Mpigi	41	407,790	159	Central 1	467,198	89,702	**14,263**
42	Mubende	177	689,530	129	Central 2	789,982	151,677	**19,566**
43	Mukono	104	795,393	129	Central 2	911,267	174,963	**22,570**
44	Nakapiripirit	61	154,494	174	North/ Karamoja Region	177,001	33,984	**5,913**
45	Nebbi	87	435,360	177	West Nile/North West	498,784	95,767	**16,951**
46	Ntungamo	52	379,987	181	South Western	435,344	83,586	**15,129**
47	Pader	139	326,338	200	North/IDP Area	373,880	71,785	**14,357**
48	Pallisa	65	520,578	116	Eastern	596,417	114,512	**13,283**
49	Rakai	111	470,365	159	Central 1	538,889	103,467	**16,451**
50	Rukungiri	36	275,162	181	South Western	315,248	60,528	**10,956**
51	Sembabule	66	180,045	159	Central 1	206,274	39,605	**6,297**
52	Sironko	26	283,092	116	Eastern	324,333	62,272	**7,224**
53	Soroti	34	369,789	116	Eastern	423,661	81,343	**9,436**
54	Tororo	78	536,888	116	Eastern	615,103	118,100	**13,700**
55	Wakiso	63	907,988	159	Central 1	1,040,265	199,731	**31,757**
56	Yumbe	16	251,784	177	West Nile/North West	288,464	55,385	**9,803**
	Total	**4,548**	**24,442,084**	**137**		28,002,853	5,376,548	**736,587**