METROPOLIS-HASTINGS BAYESIAN POISSON REGRESSION ANALYSIS WITH APPLICATION TO THE NATALITY OF MOTHERS IN LAGOS METROPOLIS

[*1,2]Ikegwu, Emmanuel M., [2]Ogundeji, Rotimi. K.
[1]Department of Statistics, Yaba College of Technology, Yaba Lagos
[2]Department of Statistics, University of Lagos, Akoka Lagos
[*]Correspondent Author: emmanuel.ikegwu@yabatech.edu.ng; +2348066940814

Abstract

This study modelled the natality of mothers in the Lagos metropolis of Lagos State using the Bayesian Poisson Regression Analysis with the Metropolis-Hastings algorithm to sample the expected posterior mean natality. The specific objectives were to compare the natality of mothers by different predictors incorporating the prior knowledge about natality with a Poisson distributed likelihood to obtain the posterior distribution. The study used nine different categorical predictors to model the natality of mothers vis mother's age, highest education qualification, religious affiliation, residence, use of contraceptives in between births, length of breastfeeding babies, length of child spacing (birth gaps), mother age at first marriage, and the Local government of residence. The prior distribution used was the normal prior on a Poisson likelihood and obtained the posterior distributions. The data used comprised 2000 mothers selected purposively and was extracted from Abe (2013), a city-wide study on infant mortality in the presence of child spacing and migration and the data were analysed using the Bayesian Poisson Regression with the help of code written in R programme environment. The study found that the expected natality of mothers in the Lagos metropolis is 2.68 (95% CI 2.46 – 2.79). Also, it found that the highest educational qualification, child spacing (birth gaps), age at first marriage and Local Government of residence has a positive impact on the natality of mothers in Lagos metropolis and while mothers' age, residence, religious affiliation, use of contraceptives in between birth, and breastfeeding length have a positive impact on natality of mothers. Also, it found that the highest educational qualification, child spacing (birth gaps) and LGA of residence have a significant impact on natality while the other predictors do not. The study therefore concludes that the Bayesian Poisson Regression Model was a good model for the natality of mothers in Lagos metropolis using the Metropolis-Hastings algorithm. It also concluded that the model determined that the expected natality of mothers falls around 3 children.

Keywords: Bayesian Poisson Regression (BPR), Birth gaps, Credible intervals, Metropolis-Hasting, Natality, Posterior mean

1.0 Introduction

The application of Bayesian analysis in virtually every field of human endeavours and demographic analysis has gained traction in recent times. Natality, or frequency of births in a community, is a critical demographic indicator that reflects the reproductive behaviour and population dynamics of a community. Understanding the factors that influence natality is essential for effective population planning, healthcare provision, and social policy formulation. Bayesian analysis has emerged as a powerful tool for modelling and predicting natality, offering a flexible framework that integrates prior knowledge with observed data to quantify uncertainty and improve decision-making.

ISSN NUMBER: 1116-249X

Bayesian methods provide a robust approach for analysing complex relationships between natality and various predictors, such as maternal age, child spacing, and contraceptive use. Unlike traditional frequentist methods, Bayesian analysis allows for the incorporation of prior information and produces a full probability distribution of the parameters, which enhances the interpretability and reliability of the results. According to van de Schoot, et al., (2021), Bayesian statistics is an approach to analyse data and estimate parameters based on Bayes' theorem. Cappello, Kim, & Palacios (2023) applied Bayesian analysis to dependent population dynamics in Coalescent models, and Ince, Paton, Kay, & Schyns, (2021) used Bayesian inference to predict population prevalence. Bayesian analysis offers several advantages over traditional statistical methods. It allows for the integration of prior information—derived from previous studies or expert knowledge—into the modelling process, which can enhance the accuracy of predictions, especially when dealing with limited or noisy data (Gelman et al., 2013) and also provides a probabilistic framework that quantifies uncertainty around parameter estimates, offering a more nuanced understanding of the factors influencing birth rates (McElreath, 2020).

Recent developments in Bayesian analysis have significantly advanced its application in demographic studies. Bayesian hierarchical models have been employed to account for variability at multiple levels, such as individual, regional, and national scales, thereby improving the precision of predictions and policy recommendations (Gelman & Hill, 2020). Additionally, the integration of Bayesian methods with modern computational techniques, such as Markov Chain Monte Carlo (MCMC) algorithms, has facilitated the handling of large and complex datasets, making Bayesian analysis more accessible and practical for researchers (Stan Development Team, 2023). One notable application of Bayesian analysis in natality research is the modelling of fertility rates using Poisson regression frameworks. These models have been used to examine the effects of maternal age, child spacing, and contraceptive use on birth rates, providing insights into how these factors interact and influence reproductive outcomes (Cameron & Trivedi, 2022). Recent studies have utilized Bayesian Poisson models to explore the impact of family planning interventions and changes in reproductive behaviour, offering valuable information for designing effective public health strategies (Cleland & Bernstein, 2021).

Moreover, Bayesian analysis has been instrumental in understanding the impact of socio-economic factors on natality. For example, Bayesian models have been used to investigate how economic conditions, education levels, and healthcare access affect birth rates, providing a comprehensive view of the socio-economic determinants of fertility (Rabe-Hesketh & Skrondal, 2021). These models allow for a nuanced analysis of how different factors contribute to natality, enabling policymakers to address the underlying causes of changes in birth rates. Bayesian analysis offers a sophisticated approach to studying natality in communities, leveraging prior knowledge and advanced computational techniques to provide detailed and reliable insights. By applying Bayesian methods, researchers can better understand the complex relationships between natality and its predictors, ultimately supporting more informed decision-making and effective policy development. Bayesian analysis has emerged as a powerful tool for understanding natality trends in communities. Its flexibility, ability to quantify uncertainty and incorporation of prior knowledge make it an attractive approach for demographic analysis.

Natality, or birth rates, is a crucial demographic indicator of a community's health, economic development, and social well-being (Hobcraft, 2003). Accurate modelling and analysis of natality trends are essential for informed decision-making and policy development. Bayesian analysis has

ISSN NUMBER: 1116-249X

emerged as a robust and flexible approach to understanding natality patterns, offering advantages over traditional statistical methods (Gelman et al., 2013). One of the earliest applications of Bayesian analysis used a Bayesian hierarchical model to estimate mortality rates in demography (Lee, 1993). This approach was later adapted to model natality rates (Brass, 1996), demonstrating the potential of Bayesian methods in demographic analysis.

Bayesian models have also been employed to analyse natality trends in various contexts. The Bayesian Poisson regression model was used to examine the effects of socioeconomic factors on birth rates in the United States (Congdon, 2001), Bayesian generalised linear model was employed to study the relationship between natality and urbanisation in Brazil (Schmertmann, Potter, & Cavenaghi, 2010), Bayesian spatio-temporal model was used to examine the impact of environmental factors on birth rates in China (Wang, Guo, & Li, 2019) and Bayesian machine learning approach was employed to predict natality trends in the United States (Guo, Li, & Wang, 2020). The Bayesian model was also developed for estimating fertility rates in developing countries (Alkema, Raftery, & Gerland, 2012), used to forecast population growth and natality rates (Raftery, et al., 2014), Bayesian spatiotemporal model to analyse the impact of socioeconomic factors on birth rates in Italy (Camarda, Raftery, & Alkema, 2017) and used Bayesian logistic regression was used to analyse the relationship between maternal education and birth outcomes in a community (Deville et al., 2014).

Bayesian analysis has also been used to address specific challenges in natality research, such as handling missing data through imputation and model averaging (Congdon, 2001) and accounting for spatial autocorrelation through capturing spatial dependencies in natality rates (Wang, Guo, & Li, 2019). Bayesian methods have been applied to various aspects of natality research, including fertility analysis, the impact of socioeconomic factors on fertility and forecast fertility trends (Schmertmann, Potter, & Cavenaghi, 2010; Raftery, 2014). Bayesian analysis has also been applied to study the determinants of birth spacing (Kozloski, Hall, & Burns, 2017) and to model the effects of birth spacing on child health outcomes (McGrath, Peterso, & Kelley, 2019). Bayesian methods have been used to analyse the risk factors of teenage pregnancy (Hobcraft, 2003) and evaluate the effectiveness of interventions aimed at reducing teenage pregnancy rates (Santelli, Rochat, & Haws, 2017). Lastly, Bayesian analysis has been applied to study the determinants of low birth weight (Camarda, Raftery, & Alkema, 2017), model the effects of low birth weight on child health outcomes (Wang, Guo, & Li, 2019) and a Bayesian hierarchical model was developed to estimate fertility rates across multiple countries, incorporating time trends and allowing for cross-country comparisons which provided robust estimates even for countries with incomplete or inconsistent data, demonstrating the utility of Bayesian methods in global demographic studies (Alkema, et al., 2011).

Also, Ruthworth, Tunniclife, & Ghosh (2017) and Wakefield (2007) used Bayesian disease mapping techniques to model birth rates in rural areas of the United Kingdom and demonstrated the utility of Bayesian spatial models in capturing local variations in birth rates, which are often masked in aggregate-level analyses. Bayesian hierarchical models were utilised to assess the impact of economic inequalities on natality rates across different Indian states which revealed significant disparities in birth rates linked to socio-economic status, highlighting the importance of addressing economic factors in public health interventions (Ghosh & Rao, 2020; Rue, Martino, & Chopin, 2009).

ISSN NUMBER: 1116-249X

Several recent studies have applied Bayesian methods to natality data, demonstrating the effectiveness of these models. Bayesian hierarchical models were used to estimate global fertility rates, accounting for variability across countries and over time (Alkema & Chao, 2022), Bayesian spatial-temporal models were explored to analyse birth rates in different regions, providing insights into spatial patterns and temporal trends in natality (Paciorek, & Liu, 2021), advancements in MCMC methods and their applications in demographic studies, including the estimation of birth rates were discussed (Roberts & Rosenthal, 2021), Applied Bayesian models was used to assess the impact of socio-economic factors on birth outcomes in Switzerland, illustrating how hierarchical models can handle complex data structures (Furrer & Helbling, 2020) and Bayesian data analysis, with applications in public health, emphasising the use of hierarchical models in demographic research was overviewed (Gelman & Vehtari, 2021).

Despite the advances in Bayesian analysis for natality research, challenges remain, including high data quality and availability which are limited in developing countries or for specific subpopulations, can be computationally intensive and require careful specification to avoid overfitting or underfitting and the results require careful interpretation and communication to ensure that findings are accurately conveyed to policymakers and practitioners, the selection of appropriate prior distributions, especially in studies where prior information is limited or subjective as incorrect priors can lead to biased results, underscoring the need for sensitivity analyses to assess the robustness of findings to different prior assumptions ; (Carlin & Louis, 2009; Gelman, et al., 2013; Deville, et al., 2014).

Bayesian analysis in research will include integrating multiple data sources such as administrative records and survey data, to improve the accuracy and completeness of natality data, developing more flexible models that can accommodate complex relationships and non-linear effects, such as machine learning algorithms, and improving communication and dissemination strategies to convey Bayesian results to policymakers and practitioners.

This study seeks to apply a Bayesian Analysis model to estimating and predicting the natality of mothers in Lagos Metropolis.

2.0 Methodology

This study adapted the model of the Bayesian Poisson Regression used by Tomal, Khan, & Wahed (2022) in modelling the natality of women in the Lagos Metropolis. In the analysis of natality data, Poisson regression comes in handy as it is commonly used to model count data due to its suitability. The predictor variables for this study include the woman's age, the woman's age at marriage, woman's education, family income level, ethnic group, religion, and use of contraceptives which other studies have shown to be predictors of the number of children born (Das, Das, & Basu Roy, 2023; Cherie, et al., 2023; Bhandari, et al., 2023; Rahman, Hossain, Rahman, & Kabir, 2022; Tomal, Khan, & Wahed, 2022; Ibeji, Zewotir, North, & Amusa, 2021; Kiser & Hossain, 2019; Upadhyay & Bhandari, 2017)

However, a Bayesian Poisson regression framework is advantageous for integrating uncertainty and prior information in modelling (Ibeji, Zewotir, North, & Amusa, 2020). The use of Markov Chain Monte Carlo (MCMC) methods for sampling from these distributions is necessary for the analytical solutions for the posterior distributions of the model parameters which are characteristically inflexible. The Metropolis-Hastings (MH) algorithm, a cornerstone of MCMC

ISSN NUMBER: 1116-249X

techniques, simplifies this sampling process, aiding robust Bayesian inference for Poisson regression models.

The data used in this study was obtained from the PhD work of the Late Dr J. B. Abe at the Obafemi Awolowo University, Ile Ife. Analysis of the data was done with the aid of Posit.

2.1 The Model (Poisson Regression Model) and The Likelihood Function

Poisson regression is personalised for modelling count data where the response variable y_i represents the number of events (natality) occurring in a fixed period or space. The model assumes that the counts are Poisson distributed. Let Y be the natality (number of Children born) of women in Lagos Metropolis in 2022 and assuming the conditional distribution of Y given the observed vector of explanatory variables **X** follows a Poisson distribution.

$$y \sim Poisson\ (\lambda) \tag{1}$$

$$P\big(Y = y \big| \boldsymbol{X}, \lambda(\boldsymbol{X};\ \boldsymbol{\beta})\big) = \frac{e^{-\ \lambda(\boldsymbol{X};\,\boldsymbol{\beta})}\ \lambda(\boldsymbol{X};\ \boldsymbol{\beta})^y}{y!}, foor\ y = 0, 1, 2, 3, \dots \tag{2}$$

Where $\lambda(\boldsymbol{X};\ \boldsymbol{\beta})$ is the conditional mean and variance of Y given **X**. The Poisson regression uses a log-link to connect the linear predictor to the mean function. Hence, we have

$$ln[E(Y|\boldsymbol{X}, \boldsymbol{\beta})] = \ln\big(\lambda(\boldsymbol{X};\ \boldsymbol{\beta})\big) = \ \boldsymbol{X}^T\boldsymbol{\beta} \tag{3}$$

Where $\boldsymbol{\beta} = (\ \beta_0,\ \beta_1\ \ \beta_2, \dots\ \ \beta_p)^{\boldsymbol{T}}$ and $\boldsymbol{X} = (1,\ x_1\ \ x_2, \dots\ \ x_p)^{\boldsymbol{T}}$ are $(p + 1)$

2.2 The Bayesian Framework

All model parameters are treated as random variables with specified prior distributions in the Bayesian approach to derive the posterior distribution of these parameters given the likelihood of the observed data. Hence, the relationship is given as

$$f(\lambda|y_1, y_2, \dots, y_n) = \frac{f(y_1, y_2, \dots, y_n|\beta)f(\lambda)}{f(y_1, y_2, \dots, y_n)} \tag{4}$$

Where $f(\lambda|y_1, y_2, \dots, y_n)$ is the posterior distribution, $f(\theta)$ is the prior distribution and $f(y_1, y_2, \dots, y_n|\lambda)$ is the likelihood function, while $\lambda = (\ \beta_0,\ \beta_1\ \ \beta_2, \dots\ \ \beta_p,\ \ \mu_1, \dots,\ \mu_j,\ \sigma_u^2, \sigma_\beta^2)$ is the vector of the model parameters. The joint posterior distribution encompasses fixed effects and hyperparameters.

2.3 The Prior Distribution

Tomal, Khan, & Wahed (2022) employed three different priors namely – the Normal, Laplace and Cauchy priors. An informative prior which reflects current knowledge and uncertainty of the parameter of interest is used in modelling.

If we assume independence among the regression coefficients $\beta_i(\boldsymbol{i} = \boldsymbol{1}, \boldsymbol{2}, \dots, \boldsymbol{p})$, the joint prior distribution for β is expressed as

ISSN NUMBER: 1116-249X

$$\pi(\boldsymbol{\beta}) = \prod_{i=1}^{p} \pi(\boldsymbol{\beta}_i|\boldsymbol{\mu_i}, \boldsymbol{\sigma_i}) \tag{5}$$

Where $\pi(\boldsymbol{\beta}_i|\boldsymbol{\mu_i}, \boldsymbol{\sigma_i})$ is the marginal prior for the ith coefficient. For $\pi(\boldsymbol{\beta}_i|\boldsymbol{\mu_i}, \boldsymbol{\sigma_i})$, three different types of priors are considered.

The Normal prior is given as

$$\pi(\boldsymbol{\beta}_i|\boldsymbol{\mu_i}, \boldsymbol{\sigma_i}) = \frac{1}{\sqrt{2\pi}\boldsymbol{\sigma_i}} exp\left\{-\frac{1}{2}\left(\frac{\boldsymbol{\beta}_i - \boldsymbol{\mu}_i}{\boldsymbol{\sigma_i}}\right)\right\} \tag{6}$$

Where the location and scale parameters are $\boldsymbol{\mu_i}$ and $\boldsymbol{\sigma_i}$.

2.4 The Posterior

The joint posterior distribution of $\boldsymbol{\beta}$ is obtained by a combination of the likelihood function and the prior distribution as follows:

$$\pi(\boldsymbol{\beta}|\boldsymbol{y}, \boldsymbol{X}) \propto L(\boldsymbol{\beta}) \times \pi(\boldsymbol{\beta}) = \prod_{i=1}^{n} P\big(\boldsymbol{Y_i} = \boldsymbol{y_i}\big|\boldsymbol{x_i}, \lambda(\boldsymbol{X};\ \boldsymbol{\beta})\big) \times \prod_{i=1}^{n} \pi(\boldsymbol{\beta}_i|\boldsymbol{\mu_i}, \boldsymbol{\sigma_i}\,) \tag{7}$$

Where L(.) and π(.) are the likelihood function and the prior distribution respectively. The joint posterior distribution of **β** does not have a closed-form expression of a standard statistical distribution. Therefore, the Monte Carlo Markov chain (MCMC) and the Gibbs sampler could not be used to generate samples from the posterior, the study adopted the Metropolis Hasting algorithm for the sampler.

2.5 The Metropolis-Hastings Algorithm

The Metropolis-Hastings (MH) algorithm is an MCMC method used to generate samples from the posterior distribution $p(\theta|y)$ when direct sampling is impracticable.

2.5.1 Target Posterior

The target is to sample from the posterior distribution:

$$p(\lambda|y) \propto f(y|\lambda)\pi(\lambda) \tag{8}$$

2.5.2 Proposal distribution

A proposal distribution $q\left(\theta^*|\ \theta^{(t-1)}\right)$ is employed to propose new parameter values θ^* based on the current state $\theta^{(t-1)}$ with a mutual choice called the Gaussian random walk given as:

$$\theta^* = \theta^{(t-1)} + \epsilon, where\ \epsilon \sim N(0, \Sigma) \tag{9}$$

Where Σ is the variance-covariance matrix of control steps.

2.5.3 Acceptance Probability

The acceptance probability α determines whether to accept the proposed move or not and is given as:

ISSN NUMBER: 1116-249X

$$\alpha(\theta,\ \theta^*|y)\ = min\left(1,\qquad \frac{p(\theta^*|y)\ q\left(\theta^{(t-1)}|\theta^*\right)}{p(\theta^{(t-1)}|y)\ q(\theta^*|\theta^{(t-1)})}\right) \tag{10}$$

When using a candidate density q for which $q\left(\theta^*|\ \theta^{(t-1)}\right) \neq q\left(\theta^{(t-1)}|\ \theta^*\right)$, acceptance ratio becomes

$$\alpha(\theta,\ \theta^*|y)\ = \frac{h(\theta^*)\ q\left(\theta^{(t-1)}|\ \theta^*\right)}{h(\theta^{(t-1)})\ q(\theta^*|\ \theta^{(t-1)})} \tag{11}$$

2.6 Metropolis-Hasting Algorithm

Initialise: Choose the initial parameter $\theta^{(0)}$

Iterate: For each iteration $t-1$ to T: Generate $\theta^* \sim q\left(\theta^*|\ \theta^{(t-1)}\right)$' Compute the ratio $r = h(\theta^*)/h\left(\theta^{(t-1)}\right) = exp\left[log\ h(\theta^*) - logh\left(\theta^{(t-1)}\right)\right]$' Compute the acceptance ratio $\alpha(\theta,\ \theta^*|y)$

Accept or reject by drawing $r \sim Uniform(0,1)$. If $r < \alpha$ set $\theta^{(t)} = \begin{cases}\theta^*, with\ probability\ r\\ \theta^{(t-1)}, with\ probability\ 1-r\end{cases}$. Under mild conditions as noted above, the draw $\theta^{(t)}$ converges in distribution to a draw from the true posterior density $p(\theta|y)$ as $t \to \infty$.

The data used in the study comprised 2000 datasets extracted from Abe (2013) which was a PhD city-wide study of the relationship between infant mortality and child spacing among Migrants and non-migrants in Lagos State. The original data was cleansed and 1566 observations were used in the study. The data was analysed with the help of code written in R programme environment.

2.7 Computations

The data with the nine (9) categorical predictors was brought into the R platform from SPSS as detailed in the code. The response variable natality (number of children born) for each woman was also included in the data set and assigned in the code appropriately and shown to follow a Poisson distribution as shown in Fig. 1. Some random true coefficients were assumed for these predictors. The total number of observations was 1566 after cleaning from a dataset of 2000 observations. The design matrix was also created while the prior assumed was a normal distribution prior, the likelihood was already declared a Poisson distribution and the posterior was generated as a normal distribution giving a normal-normal distribution. The Metropolis-Hastings algorithm to sample from the posterior distributions of the regression coefficients was executed with a sample of 20,000 iterations and 500 burn-ins at 0.5 steps. Lastly, the posterior means were computed, summarised and interpreted.

ISSN NUMBER: 1116-249X

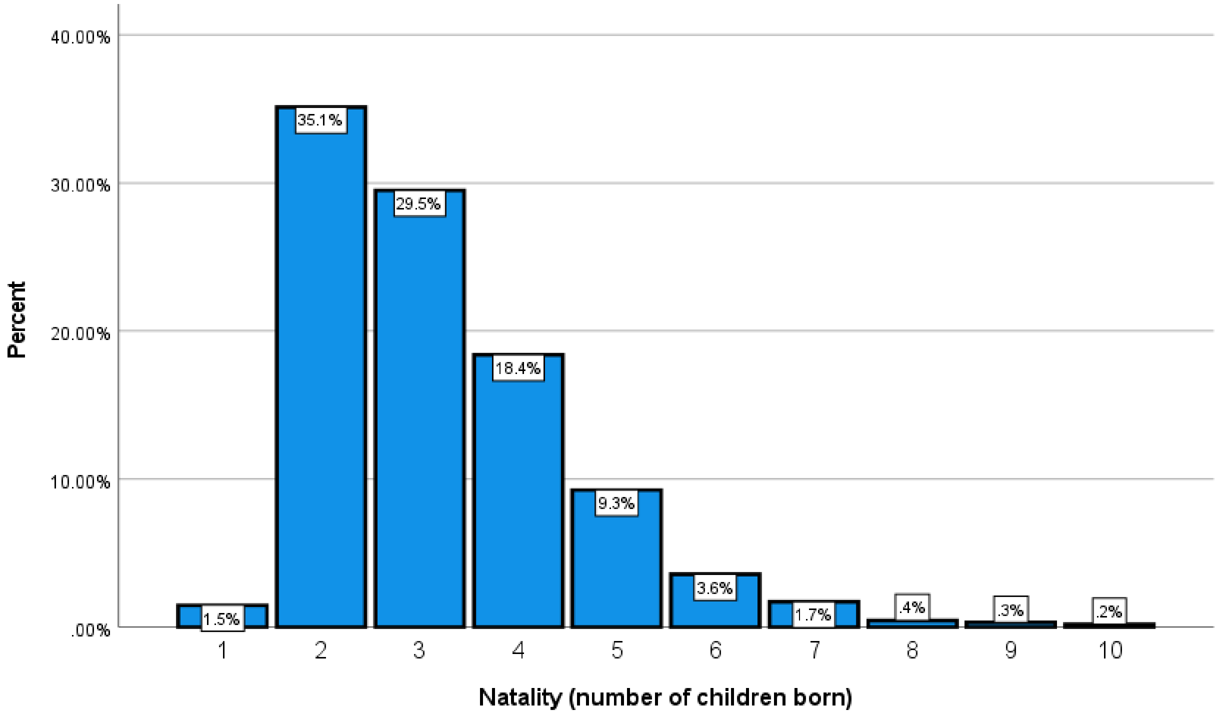

Fig. 1: The bar chart of the distribution natality (number of children born) of mothers in Lagos metropolis

The chart distribution shows that the data is suitable for Poisson regression as it looks a lot like a Poisson distribution. The predictor variables included in the study include place of residence (urban or rural), age of mothers (15 – 34 years, 35 – 54 years and 55 and above years), religion of mother (Christianity and Islam), Highest educational qualification of mother (No Formal education, Primary education, Secondary education, Other higher school, University), contraceptive use in between births (Yes or No), Length of breastfeeding baby (<= 6 months, <= 12 months, <= 18 months, <= 24 months, Above 24 months), Child spacing length (<= 12 months, <= 18 months, <= 24 months, <=30 months, Above 30 months), Age at first marriage (<= 20 years, 21 – 30 years, Above 30 years) and 10 selected Local Government Areas of the State.

3.0 Results

The result of the data collected from women in Metropolitan Lagos selected from 10 Local Government Areas and analysed using the Bayesian Poisson Regression Analysis with Metropolis-Hasting Algorithm.

ISSN NUMBER: 1116-249X

Table 1: Descriptive Statistics of the predictor variables in the study

Variables	Category	Frequency	Percent
Age of mothers (years)	15 – 34	655	41.8
	35 – 54	781	49.9
	55 – 74	130	8.3
Highest level of Education	None	23	1.5
	Primary level	209	13.3
	Secondary level	568	36.3
	Other higher school	324	20.7
	University	442	28.2
Religious affiliation	Christianity	1119	71.5
	Islam	447	28.5
	Traditional	-	-
Place of residence	Rural	695	44.4
	Urban	871	55.6
Age at first marriage	<= 20 years	388	24.8
	21 - 30 years	1142	72.9
	Above 30 years	36	2.3
Contraceptive use in between births	Yes	521	33.3
	No	1045	66.7
Breastfeeding length before sexual intercourse	6 months	907	57.9
	12months	445	28.4
	18 months	111	7.1
	24 months	83	5.3
	Above 24 months	20	1.3
Child spacing length	12 months	181	11.6
	18 months	343	21.9
	24 months	728	46.5
	30 months	194	12.4
	Above 30 months	120	7.7
Local Government Area of residence	Surulere	139	8.9
	Eti-Osa	140	8.9
	Ajeromi-Ifelodun	181	11.6
	Ikeja	150	9.6
	Lagos Mainland	156	10.0
	Ojo	160	10.2
	Ikorodu	164	10.5
	Alimosho	145	9.3
	Badagry	156	10.0
	Epe	175	11.2

All the predictor variables are categorical hence, their frequency distribution is given.

ISSN NUMBER: 1116-249X

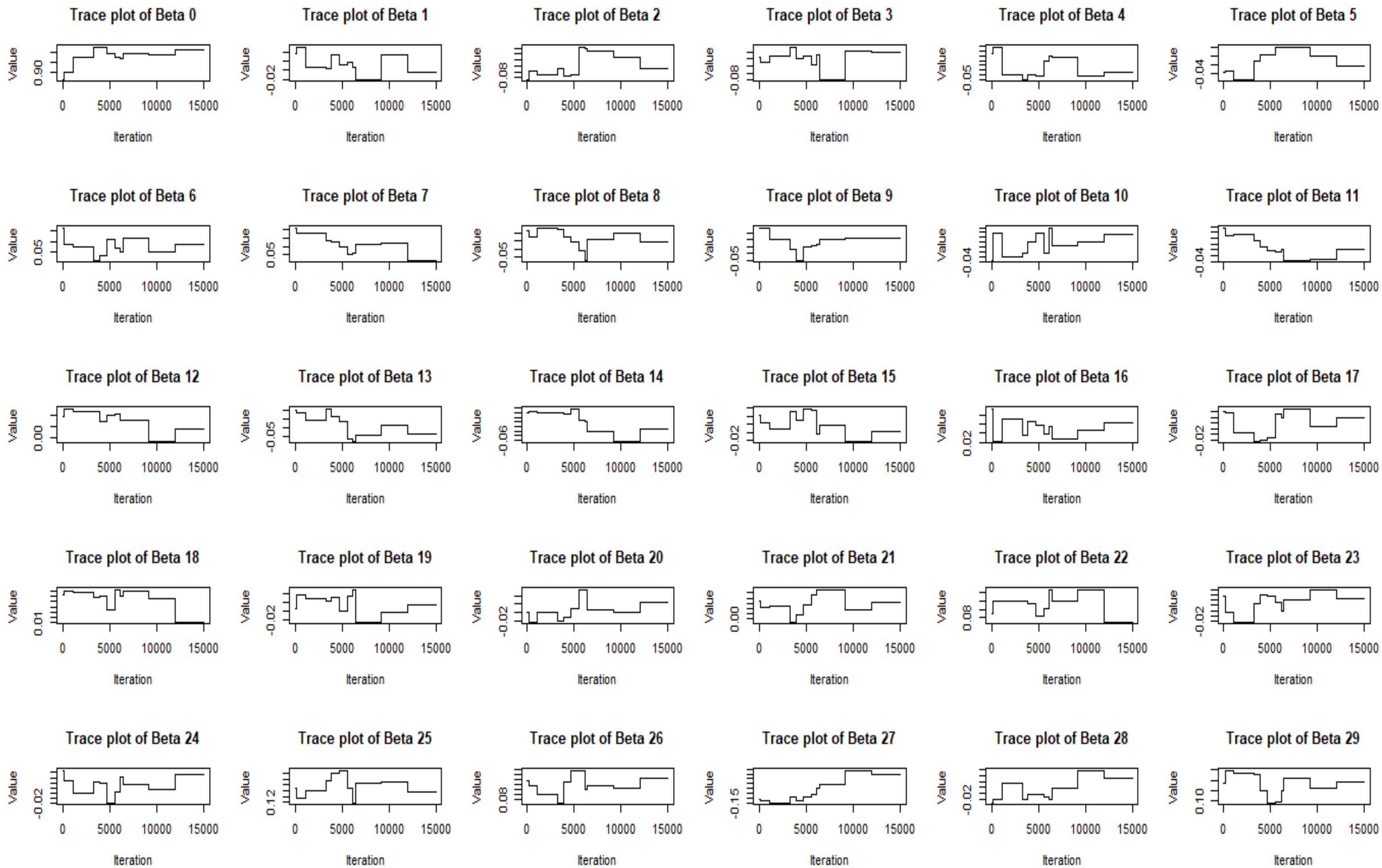

Fig. 2: The trace plot of the simulated predictor variables in the study

ISSN NUMBER: 1116-249X

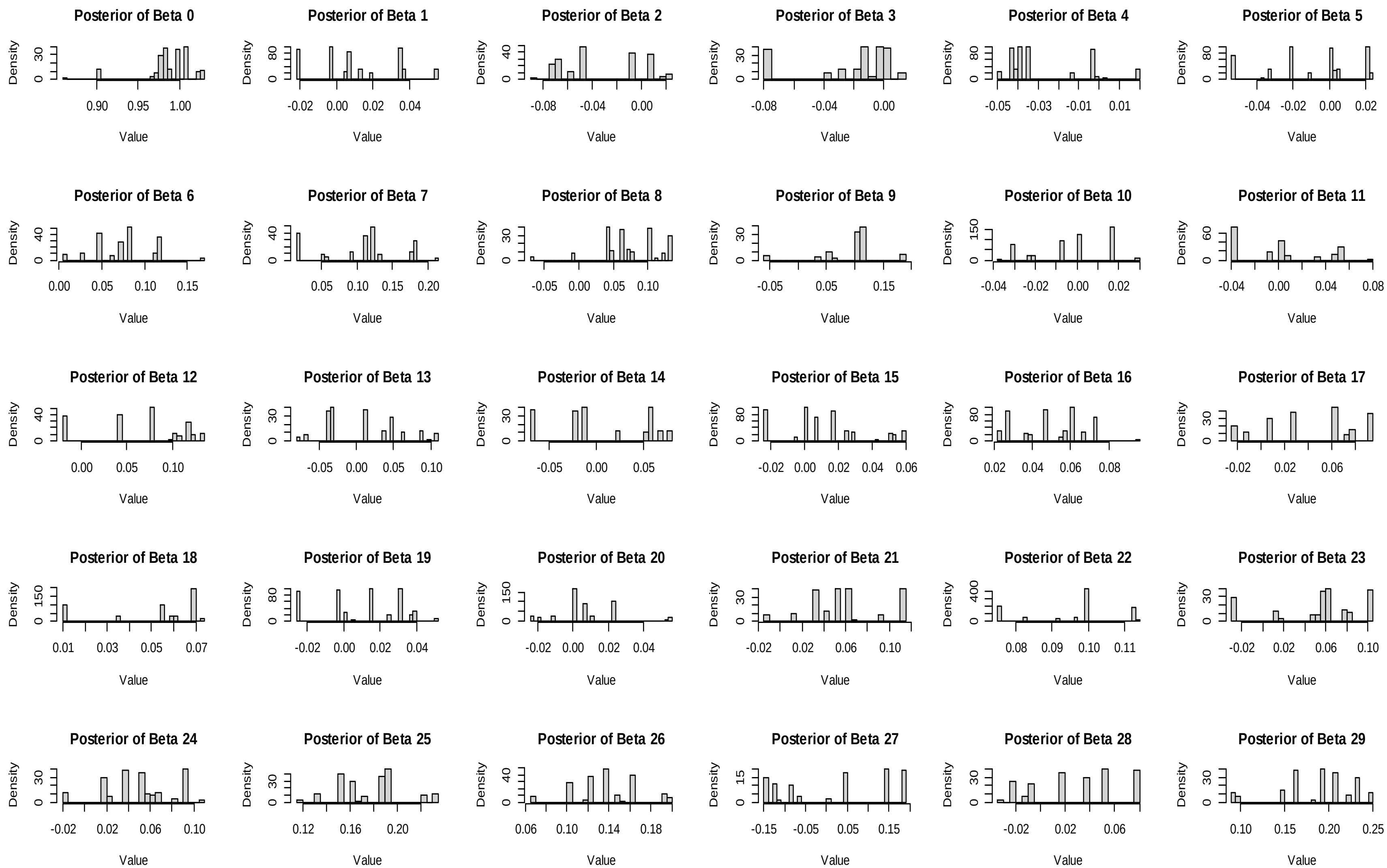

Fig. 3: The histogram of the simulated predictor variables in the study

ISSN NUMBER: 1116-249X

Table 2: Bayesian Poisson Regression Model Summary with 95% Credible Intervals

Variables	Categories	Posterior Mean	95% Credible Intervals	
			CI Lower	**CI Upper**
Intercept		0.987	0.901	1.026
Residence	Rural	1	-	-
	Urban	0.011	-0.021	0.054
Religion	Christianity	1	-	-
	Islam	-0.033	-0.074	0.024
	Traditional	-0.020	-0.078	0.015
Age (years)	15 - 34	1	-	-
	35 - 54	-0.027	-0.049	0.019
	55 & above	-0.009	-0.053	0.022
Highest educational qualification	None	1	-	-
	Primary level	0.107	0.016	0.181
	Secondary level	0.075	-0.008	0.135
	Other higher school	0.076	0.007	0.115
	University	0.095	-0.050	0.183
Contraceptive Use	No	1	-	-
	Yes	-0.001	-0.031	0.018
Breastfeeding length	6 months	1	-	-
	12 months	0.008	-0.067	0.109
	18 months	-0.0004	-0.037	0.051
	24 months	0.067	-0.017	0.134
	Above 24 months	0.005	-0.065	0.078
Birth gaps (child spacing)	12 months	1	-	-
	18 months	0.011	-0.022	0.059
	24 months	0.050	0.022	0.072
	30 months	0.042	-0.025	0.093
	Above 30 months	0.052	0.011	0.074
Age at first marriage (years)	15 – 20	1	-	-
	21 – 30	0.010	-0.025	0.039
	Above 30	0.007	-0.023	0.055
LGA	Ajeromi-Ifelodun	1	-	-
	Alimosho	0.060	-0.012	0.114
	Badagry	0.096	0.075	0.112
	Epe	0.054	-0.025	0.103
	Eti-Osa	0.050	-0.018	0.092
	Ikeja	0.177	0.135	0.232
	Ikorodu	0.137	0.067	0.195
	Lagos Mainland	0.025	-0.145	0.188
	Ojo	0.029	-0.023	0.079
	Surulere	0.188	0.092	0.248

ISSN NUMBER: 1116-249X

Table 1 shows the posterior means for each of the regression coefficients, which are the most likely values for the coefficients given the prior information and the data obtained and used for the Bayesian Poisson Regression (BPR) with the Metropolis-Hastings Algorithm and their 95% credible intervals. Nine categorical predictors were employed to predict the expected Natality (number of children born) by Mothers in the Metropolis of Lagos State.

The table shows that living in the urban area increases the expected count of natality (0.011) but its effect is not significant with credible intervals including zero (-0.021, 0.054). This implies that urban dwellers have 1.1% more natality than rural residents (95% CI: -2.1% - 5.4%) showing that the expected natality of urban residents is 2.7 with a 95% credible band of 2.4 to 2.9.

Also, it shows that religion decreases the expected natality count of Lagos mothers but not significantly with Islam (-0.033, CI: -0.074 – 0.024) and traditional religion (-0.020, CI: -0.078 – 0.015). This shows that the expected natality count of Moslems in Lagos metropolis is 3.3% less than that of Christians (95% CI is -7.4% to 2.4%) which gives expected natality of 2.6 with a 95% credible band of 2.27 to 2.86 while that of Traditionalists is 2.0% less than that of Christians (95% CI is -7.8% to 1.5%) giving an expected natality of 2.6 and 95% credible band of 2.28 to 2.83.

In the same vein, mothers age also decreases the expected natality of the women but not significantly with 25 – 55 years (-0.027, CI: -0.049 – 0.018) and 55 years and above (-0.009, CI: -0.053 – 0.022) implying that mothers aged 25 – 55 years have 2.7% less expected natality (95% CI -4.9 – 1.8%) than those 15 – 14 years. Hence, mothers 25 – 55 years have an expected natality count of 2.61 with a 95% credible band of 2.34 – 2.84. Likewise, mothers aged 55 years and above have 0.9% less expected natality (95% CI -5.3 – 2.2%) than those 15 – 14 years. Therefore, mothers 55 years and above have an expected natality count of 2.66 with a 95% credible band of 2.33 – 2.85.

In addition, the table shows that the highest educational qualification positively adds to the expected count of natality with primary education (0.107, CI: 0.016 – 0.181) and other higher schools (0.076, CI: 0.007 – 0.115) having a significant impact while secondary level (0.075, CI: -0.008 – 0.135) and university education (0.095, CI: -0.050 – 0.183) do not have a significant impact. The implication is that mothers with primary education have 10.7% more expected natality (95% CI 1.6 – 18.1%) than mothers with no formal education and with expected natality of 2.99 and 95% credible band of 2.50 – 3.34. Also, mothers with secondary education have 7.5% more expected natality (95% CI 0.8 – 13.5%) than mothers with no formal education, and with expected natality of 2.89 and 95% credible band of 2.44 – 3.19. In the same vein, mothers with other higher schools have 7.6% more expected natality (95% CI 0.7 – 11.5%) than mothers with no formal education, and with expected natality of 2.90 and 95% credible band of 2.48 – 3.13. Lastly, mothers with other higher schools have 9.5% more expected natality (95% CI -5.0 – 18.3%) than mothers with no formal education, and with expected natality of 2.44 and 95% credible band of 2.34 – 3.35.

Also. the use of contraceptives before the next pregnancy negatively impacts the natality of mothers but not significantly (-0.001, CI: -0.031 - 0.018). This shows that mothers who use contraceptives in between births have 0.1%% more expected natality (95% CI -3.1 – 1.8%) than mothers who do not use contraceptives and with expected natality of 2.68 and 95% credible band of 2.39 – 2.84.

Furthermore, the length of time the mother breastfed their children has mixed effect on the expected count natality of mothers with 18 months (-0.0004, CI: -0.037 – 0.051) exhibiting

ISSN NUMBER: 1116-249X

negative impact and 12 months (0.008, CI: -0.067 – 0.109), 24 months (0.067, CI: -0.017 – 0.134) and above 24 months (0.005, CI: -0.065 – 0.078) having positive impact on the expected count but none of them is significant. This shows that mothers who breastfed their children 12 months have 0.8% more expected natality (95% CI -6.7 – 10.9%) than mothers who breastfed their babies only 6 months and with expected natality of 2.70 and 95% credible band of 2.30 – 3.11. Also, mothers who breastfed their children 18 months have 0.04% less expected natality (95% CI -3.7 – 5.1%) than mothers who breastfed their babies only 6 months, and with expected natality of 2.68 and 95% credible band of 2.37 – 2.94. Similarly, mothers who breastfed their children 24 months have 6.7% more expected natality (95% CI -1.7 – 13.4%) than mothers who breastfed their babies only 6 months, and with expected natality of 2.87 and 95% credible band of 2.42 – 3.19. Lastly, mothers who breastfed their children above 24 months have 0.5% more expected natality (95% CI -6.5 – 7.8%) than mothers who breastfed their babies only 6 months, and with expected natality of 2.70 and 95% credible band of 2.31 – 3.02.

Likewise, the birth gaps (the number of months of child spacing) showed a positive influence on the expected count of natality of mothers with 18 months (0.011, CI: -0.022 – 0.58), 24 months (0.050, CI: -0.022 – 0.072), 30 months (0.042, CI: -0.025 – 0.093) and above 30 months (0.051, CI: -0.011 – 0.074) but none of the period has significant impact. This implies that mothers who have 18 months of child spacing (birth gaps) have 0.8% more expected natality (95% CI -6.7 – 10.9%) than mothers who had 6 months of child spacing and with expected natality of 2.71 and 95% credible band of 2.41 – 2.96. Also, the results show that mothers who have 24 months of child spacing (birth gaps) have 5.0% more expected natality (95% CI -2.2 – 7.2%) than mothers who had 6 months of child spacing and with expected natality of 2.82 and 95% credible band of 2.41 – 3.00. Lastly, it shows that mothers who have 30 months of child spacing (birth gaps) have 5.1% more expected natality (95% CI -1.1 – 7.4%) than mothers who had 6 months of child spacing and with expected natality of 2.82 and 95% credible band of 2.44 – 3.00.

It further shows that the mother's age at first marriage positively but not significantly impact the expected count of natality of mothers in Lagos metropolis with 21 – 30 years (0.010, CI: -0.025 – 0.039) and above 30 years (0.007, CI: -0.023 – 0.055). The results show that mothers who were 21 – 30 years at first marriage have 1.0% more expected natality (95% CI -2.5 – 3.9%) than mothers who were 20 years or below and with expected natality of 2.71 and 95% credible band of 2.40 – 2.90. In addition, it shows that mothers who were above 30 years at first marriage have 0.7% more expected natality (95% CI -2.3 – 5.5%) than mothers who were 20 years or below and with expected natality of 2.70 and 95% credible band of 2.41 – 2.95.

Lastly, the Local Government Area of residence also has a positive impact on the expected count of natality among mothers with Badagry (0.096, CI: 0.075 – 0.112), Ikeja (0.177, CI: 0.135 – 0.232), Ikorodu (0.137, CI: 0.067 – 0.195) and Surulere (0.188, CI: 0.092 – 0.248) all significant and Alimosho (0.060, CI: -0.012 – 0.114), Epe (0.054, CI: -0.025 – 0.103), Eti-Osa (0.050, CI: -0.018 – 0.092), Lagos Mainland (0.025, CI: -0.145 – 0.188) and Ojo (0.029, CI: -0.023 – 0.079) not significant. Therefore, the results show that mothers who live in Badagry have 9.6% more expected natality (95% CI 7.5 – 11.2%) than mothers who live in Ajeromi-Ifelodun, and with expected natality of 2.95 and 95% credible band of 2.65 – 3.12. Also, it shows that mothers who live in Ikeja have 17.7% more expected natality (95% CI 13.5 –23.2%) than mothers who live in Ajeromi-Ifelodun and with expected natality of 3.20 and 95% credible band of 2.82 – 3.52. In the same vein, shows that mothers who live in Ikorodu have 13.7% more expected natality (95% CI

ISSN NUMBER: 1116-249X

6.7 – 19.5%) than mothers who live in Ajeromi-Ifelodun and with expected natality of 3.08 and 95% credible band of 2.63 – 3.39. Similarly, shows that mothers who live in Surulere have 18.8% more expected natality (95% CI 9.2 – 24.8%) than mothers who live in Ajeromi-Ifelodun and with expected natality of 3.24 and 95% credible band of 2.70 – 3.58. In addition, it shows that mothers who live in Alimosho have 6.0% more expected natality (95% CI -1.2 – 11.4%) than mothers who live in Ajeromi-Ifelodun and with expected natality of 2.85 and 95% credible band of 2.43 – 3.13. Also, it shows that mothers who live in Epe have 5.4% more expected natality (95% CI -2.5 – 10.3%) than mothers who live in Ajeromi-Ifelodun and with expected natality of 2.83 and 95% credible band of 2.40 – 3.09. Furthermore, it shows that mothers who live in Eti-Osa have 5.0% more expected natality (95% CI -1.8 – 9.2%) than mothers who live in Ajeromi-Ifelodun and with expected natality of 2.85 and 95% credible band of 2.42 – 3.06. Likewise, it shows that mothers who live in Lagos Mainland have 2.5% more expected natality (95% CI -14.5 – 18.8%) than mothers who live in Ajeromi-Ifelodun and with expected natality of 2.75 and 95% credible band of 2.13 – 3.37. Lastly, it shows that mothers who live in Badagry have 2.9% more expected natality (95% CI -2.3 – 7.9%) than mothers who live in Ajeromi-Ifelodun and with expected natality of 2.76 and 95% credible band of 2.41 – 3.02.

2.4 Discussion

The study did not find any statistically significant difference in the expected natality of mothers in the rural and urban areas which agreed with Tomal, Khan, & Wahed (2022) that the gap between them is narrowing in Lagos metropolis because almost all areas in Lagos State is seen by outsiders as urban signifying the narrow sense thereof in Lagos metropolis. The study also found no significant difference in the expected natality of Lagos metropolis mothers in the three major different religious affiliations in Nigeria in Lagos metropolis. This result disagreed with Ibeji, et al, (2020) who reported that religion is significant for a generalised Poisson model which is frequentist and different from a Bayesian Poisson Regression model with the Metropolis-Hastings algorithm.

The study found a significant difference in the expected natality of mothers who attained primary education and other higher school with mothers with no formal education but no significant difference between those who attained secondary education and university with mothers with no formal education in the Lagos metropolis. These results partly agree with other studies that found that the expected natality of mothers (children ever born) is significantly associated with mothers' highest education level (Bhandari, et al., 2023; Rahman, et al., 2022; Ibeji, et al, 2020). However, the disagreement seems justified because these studies used frequentist models while this study applied a Bayesian Poisson regression model with the Metropolis-Hastings algorithm for sampling the posterior means. Though mothers who use contraceptives in between births have lower expected natality than those who do not use them, the difference is not significant in the Lagos metropolis. This agrees with Upadhyay & Bhandari (2017) who found that women who use contraceptives have fewer children as contraceptive use is a family planning method. However, the result differs in the significance of the difference which is a result of this study using the Bayesian Poisson regression model while the other used only the OLS. However, Gebre (2024) also reported that women using contraceptives have less expected natality than those who do not while using a negative binomial regression method.

In addition, the study found no significant difference in the expected natality of mothers with different lengths of breastfeeding in the Lagos metropolis. However, this finding is at variance with the report of Mekebo, et al. (2022) that women with large families (a greater number of

ISSN NUMBER: 1116-249X

children) tend to reduce the length of breastfeeding. This is inconclusive as we could not find any study that has directly considered the association or effect of breastfeeding length on the number of children by mothers. The study also found that there is a significant difference in the expected natality of mothers who had 24 months and above 30 months of child spacing (birth gaps) and mothers who spaced only 12 months but no significant difference with those who spaced 18 and 30 months in Lagos metropolis.

Furthermore, there is no significant difference in the expected natality of mothers who married at different ages in the Lagos metropolis. This negates the findings of Upadhyay & Bhandari (2017) who reported that marrying early increases the chances of a higher number of children being born. Lastly, the study found significant differences in the expected natality of mothers in the Lagos metropolis who live in different Local Government Areas. These study findings are in agreement with other studies from other jurisdictions that the expected natality of children ever-born differ significantly by region or division or metropolis and Local Government Area (Bhandari, et al., 2023; Tomal, Khan, & Wahed, 2022; Ibeji, et al., 2020).

2.5 Conclusions

The study explored a Bayesian Poisson regression model using the Metropolis-Hasting Algorithm to sample the Posterior distributions and determined the mean posterior parameters with their 95% credible intervals. The study concluded that while some of the predictors showed a significant impact on the expected natality of mothers in the Lagos metropolis, others did not. The Bayesian inculcation to estimating the posterior parameters using a normal prior applied to a Poisson likelihood had shown that the credible intervals are indeed better than the confidence intervals obtained in a frequentist analysis.

Acknowledgement: The authors do sincerely appreciate the Late Dr J. B. Abe who collected the data used for this study during his PhD work.

Conflict of Interest: The authors declare that there is no conflict of interest in the course of this study nor in its reporting.

References

Abe, J. B. (2013) *Infant Mortality and Child Spacing among Migrants and non-Migrants in Lagos State*. A PhD Thesis submitted to the Department of Demography and Social Science, Obafemi Awolowo University, Ile-Ife, Osun State.e

Alkema, L., & & Chao, F. (2022). Global fertility trends: A Bayesian perspective. *Population Studies, 76*(1), 45 - 62.

Alkema, L., Raftery, A. E., & Gerland, P. (2012). Bayesian estimation of fertility rates from incomplete data. *Journal of the Royal Statistical Society: Series A, 175*(3), 761 - 778.

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

Bhandari, K. R., Ghimire, K., Upadhyay, H. P., Prachhai, D. B., & Karki, S. (2023). Children Ever Born and its Associated Factors of Women in Tansen Municipality of Palpa District Nepal.

ISSN NUMBER: 1116-249X

International Journal of Silkroad Institute of Research and Training, 1(2), 84 - 90. doi:10.3126/ijsirt.v1i2.61665

Brass, W. (1996). Demographic data analysis in developing countries. *Journal of the Royal Statistical Society: Series A, 159*(3), 541 - 557.

Camarda, C. G., Raftery, A. E., & Alkema, L. (2017). Bayesian spatio-temporal modelling of socioeconomic factors affecting birth rates in Italy. *Journal of Applied Statistics, 44*(1), 147 - 162.

Cameron, A. C., & Trivedi, P. K. (2022). *Regression Analysis of Count Data* (3rd ed.). Cambridge: University Press.

Cappello, L., Kim, J., & Palacios, J. (2023). Bayesian Inference of Dependent Population Dynamics in Coalescent Models. *PLOS Computational Biology, 19*(3), 1010897. doi:10.1371/journal.pcbi.1010897

Carlin, B. P., & Louis, T. A. (2009). *Bayesian Methods for Data Analysis* (3rd ed.). Boca Raton: Chapman & Hall/CRC.

Cherie, N., Getacher, L., Belay, A., Gultie, T., Mekuria, A., Sileshi, S., & Degu, G. (2023). Modelling on number of children ever born and its determinants among married women of reproductive age in Ethiopia: A Poisson regression analysis. *Heliyon, 9*(3), e13948. doi:10.1016/j.heliyon.2023.e13948

Cleland, J., & Bernstein, S. (2021). The Impact of Contraceptive Use on Birth Rates: A Demographic Perspective. *Population Studies, 75*(4), 541 - 559.

Congdon, P. (2001). *Bayesian statistical modelling.* New York: John Wiley & Sons.

Das, P., Das, T., & Basu Roy, T. (2023). Role of proximate and non-proximate determinants in children ever born among Indian women: Change detection analysis from NFHS-3 & 5. *Women and Children Nursing, 1*(1), 9 - 17. doi:10.1016/j.wcn.2023.03.00

Deville, P., Song, C., Eagle, N., Blondel, V. D., Barabási, A. L., & Wang, D. (2014). Scaling identity connects human mobility and social interactions. *Proceedings of the National Academy of Sciences, 111*(49), 17480 - 17485.

Furrer, R., & Helbling, D. (2020). Bayesian hierarchical models in public health: Assessing birth outcomes in Switzerland. *Journal of the Royal Statistical Society: Series A, 183*(2), 543 - 560.

Gebre, M. N. (2024). Number of children ever-born and its associated factors among currently married Ethiopian women: Evidence from the 2019 EMDHS using negative binomial regression. *BMC Women's Health, 24*, 95. doi:10.1186/s12905-024-02883-w

Gelman, A., & Hill, J. (2020). *Data Analysis Using Regression and Multilevel/Hierarchical Model.* Cambridge: University Press.

ISSN NUMBER: 1116-249X

Gelman, A., & Vehtari, A. (2021). *Bayesian Data Analysis* (4th ed.). New York: CRC Press.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis* (3rd ed.). Boca Raton: CRC Press.

Ghosh, M., & Rao, J. N. (2020). Bayesian Hierarchical Modelling: A Review:. *Statistical Sciences, 35*(1), 15 - 32.

Guo, Y., Li, Y., & Wang, F. (2020). Bayesian machine learning for predicting natality trends in the United States. Journal of Applied Statistics. *47*(1), 147 - 162.

Hobcraft, J. (2003). Reflections on demographic transition. *Population and Development Review, 29*(3), 359 - 373.

Ibeji, J. U., Zewotir, T., North, D., & Amusa, L. (2020). Modelling fertility levels in Nigeria using Generalized Poisson regression-based approach. *Scientific African, 9*, e00494. doi:10.1016/j.sciaf.2020.e00494

Ibeji, J. U., Zewotir, T., North, D., & Amusa, L. (2021). Modelling children ever born using performance evaluation metrics: A dataset. *Data in Brief, 36*, 107077. doi:10.1016/j.dib.2021.107077

Ince, R. A., Paton, A. T., Kay, J. W., & Schyns, P. G. (2021). Bayesian inference of population prevalence. *eLife*. doi:10.7554/eLife.62461

Kiser, H., & Hossain, M. A. (2019). Estimation of number of ever born children using zero truncated count model: evidence from Bangladesh Demographic and Health Survey. *Health Information and Scientific System, 18*, 064. doi:10.1007/s13755-018-0064-y

Kozloski, M., Hall, R., & Burns, T. (2017). Bayesian Analysis of Birth Spacing and Child Health Outcomes. *Journal of Applied Statistics, 44*(7), 1185 - 1203.

Lee, R. D. (1993). Inverse projection and demographic fluctuations: A Bayesian approach. *Population Studies, 47*(1), 147 - 162.

McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan* (2nd ed.). Boca Raton: CRC Press.

McGrath, M., Peterso, D., & Kelley, C. (2019). Bayesian Modeling of Low Birth Weight: A Hierarchical Approach. *Statistical Modelling, 19*(5), 453 - 472.

Mekebo, G. G., Argawu, A. S., Likassa, H. T., Ayele, W., Wake, S. K., Bedada, D., . . . al, e. (2022). Factors influencing exclusive breastfeeding practice among under-six months infants in Ethiopia. *BMC Pregnancy and Children, 22*, 630. doi:10.1186/s12884-022-04955-x

Paciorek, C. J., & Liu, Z. (2021). Bayesian spatial-temporal modelling of birth rates: Applications in epidemiology. *Statistical Methods in Medical Research, 30*(4), 1138 - 1150.

Rabe-Hesketh, S., & Skrondal, A. (2021). *Multilevel and Longitudinal Modeling Using Stata* (3rd ed.). Stata Press.

Raferty, A. E. (2014). *Bayesian Analysis in the Social Sciences.* New York: John Wiley & Sons.

Rahman, A., Hossain, Z., Rahman, M. L., & Kabir, E. (2022). Determinants of children ever born among ever-married women in Bangladesh: Evidence from the Demographic and Health Survey 2017–2018. *BMJ Open, 12*(6), e055223. doi:10.1136/bmjopen-2021-055223

Roberts, G. O., & Rosenthal, J. S. (2021). Advances in MCMC methods for Bayesian demographic studies. *Annual Review of Statistics and Its Application, 8*, 79 - 101.

Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian Inference in Latent Gaussian Models by using Integrated Nested Laplace Approximation. *Journal of the Royal Statistical Society: Series B, 71*(2), 319 - 392.

Ruthworth, A., Tunniclife, D., & Ghosh, S. (2017). Bayesian Disease Mapping Techniques: Advances and Applications. *Statistical Modelling, 17*(3), 239 - 265.

Santelli, J. S., Rochat, R., & Haws, S. (2017). Bayesian evaluation of teenage pregnancy interventions. *Journal of Adolescent Health, 60*(4), 380 - 387.

Schmertmann, C. P., Potter, J. E., & Cavenaghi, S. M. (2010). Bayesian modelling of urbanization and fertility in Brazil. *Demography, 47*(3), 541 - 557.

Stan Development Team. (2023). *Stan User's Guide*. Retrieved August 23, 2024, from mc-stan.org/users/documentation/

Tomal, J. H., Khan, J. R., & Wahed, A. S. (2022). Weighted Bayesian Poisson Regression for The Number of Children Ever Born per Woman in Bangladesh. *Journal of Statistical Theory and Applications, 21*, 79 - 105. doi:10.1007/s44199-022-00044-2

Upadhyay, H. P., & Bhandari, K. R. (2017). Factors Associated with Children Ever Born: A Case Study of SomadiVillage Development Committee of Palpa District, Nepal. *Advanced Journal of Social Science*, 15 - 29. doi:http://dx.doi.org/10.21467/ajss.1.1.15-29

Upadhyay, H., & Bhandari, K. (2017). Factors Associated with Children Ever Born: A Case Study of Somadi Village Development Committee of Palpa District, Nepal. *Advanced Journal of Social Science, 1*(1), 15 - 29. doi:10.21467/ajss.1.1.15-29

van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., . . . Yau, C. (2021). Bayesian statistics and modelling. *Nature Reviews Methods Primers, 1*(3), 43586. doi:10.1038/s43586-020-00003-0

Wakefield, J. (2007). Disease mapping and spatial regression with count data. *Biostatistics, 8*(2), 158 - 183.

Wang, F., Guo, Y., & Li, Y. (2019). Bayesian spatio-temporal modelling of environmental effects on birth rates in China. *Journal of Exposure Science & Environmental Epidemiology, 29*(1), 34 - 42.