GEOADDITIVE MODELLING OF ADOLESCENT AGE OF SEXUAL INITIATION AMONG NIGERIAN WOMEN

[1]Abiodun, A.A and [2]Adeleke, K.A.

[1]Department of Statistics, University of Ilorin.

[2]Department of Mathematics, Obafemi Awolowo University, Ile-Ife.

Correspondence author: alfredabiodun1@gmail.com

(Received: 18th April, 2016; Accepted: 7th June, 2016)

ABSTRACT

This study investigated the geographical variations in adolescent age at sexual initiation among Nigerian women aged 15-49 years using geoadditive binary probit model. The effect of current age was assumed to be nonlinear. Geographical information at state level was also explored in the model to assess the impact of spatial distribution of adolescent sexual initiation. Effects of all categorical covariates were assumed to be linear and hence estimated in a usual parametric form. Inferences were based on Markov Chain Monte Carlo (MCMC) techniques, while model selection was based on Deviance Information Criteria (DIC). Findings revealed significant spatial variations on level of adolescent sexual initiation among Nigerian women as well as a clear evidence of nonlinear effect of age.

Keywords: Adolescent, Geoadditive, Nonparametric, Spatial-effects, Bayesian

INTRODUCTION

There have been changes over the years in the age at which adolescents initiate sex in Nigeria. Studies have examined early sexual activities largely as a potential risk factor for adverse outcomes rather than identifying the correlates of the timing of sexual debut (Pettifor *et al*., 2005; Harrison *et al*., 2005). Situation analysis on HIV and AIDS in Nigeria revealed that the HIV prevalence rate has reduced from 5.8% in 2001 (the peak) to 4.4% in 2005 (FMOH, 2006). The adolescent age at which women had their first sexual intercourse ranged from10 to19 years with mean age of 13.4±1.5 (Nnebue *et al*., 2016).

Early age at sexual initiation exposes adolescents to serious health consequences such as unwanted pregnancy; sexually transmitted infections (STIs), including Human Immunodeficiency Virus (HIV); increased risk of human papilloma virus infection; cancer of the cervix and delays in maturation toward healthy adult psychosocial adjustment. (Dickson *et al*., 1998; Ekundayo *et al*., 2007; Mmbaga *et al*., 2012).

As adolescents mature, they acquire the knowledge and skills to negotiate for sex. (Masatu, *et al*., 2009,). Therefore, delaying adolescent age at sexual initiation has been key in most interventions in sub-Saharan Africa. (WHO, 2004; Nnebue *et al*., 2016)

Studies from the United States (U.S.) suggested that adolescents who have fewer interpersonal and community resources to draw upon are likely to initiate their sexual activity earlier (Brewster, 1994; Abiodun *et al*., 2014)

Most studies in sub-Saharan Africa have focused on the social, demographic and familial factors associated with sexual initiation and reasons adolescents begin to have consensual intercourse (Andersen-Ellstrom *et al*., 1996; Kinsman *et al*., 1998 and Rink *et al*., 2007). Such studies have only reported geographical variations at a highly aggregated regional levels and trends for data collected from similar multi-round surveys (Agha *et al*., 2006 and Chiao and Mishra, 2007). A major challenge with this is that information is concealed within regions and has an adverse influence on policy formulation. First, policy is not usually formulated at a zonal level since five to seven states may constitute a zone.

In this study therefore, we have applied a geoadditive binary probit model to adolescent age at sexual initiation in Nigeria. This allows us to simultaneously assess the common influence of certain factors, which could possibly be associated with adolescent age at which Nigerian women initiated sex.

MATERIALS AND METHODS

The Study Data

The data used for this study were drawn from Nigeria Demographic and Health Survey (NDHS) for 2013 (www.measuredhs.com). The 2013 NDHS was conducted by the National Population Commission (NPC) with funding support from U.S Agency for International Development (USAID), the United Nations Population Fund (UNFPA) and the United Kingdom Department for International Development (DFID). Technical support was provided by ICF International. The 2013 NDHS sample was selected using a three-stage stratified design consisting of 904 clusters, 372 urban areas and 532 in rural areas. For the purpose of this paper, a database for females aged 15 to 49 years, was created from the main data. Information about their adolescent sexual behaviors were obtained.

Adolescent Age at Sexual Initiation

The data consists of 31,482 women aged 15-49 years. The response variable was the adolescent age at which a respondent initiated sex. Out of this number, 7991 (25.4%) had initiated sex during their adolescent age. Adolescent age was categorized into early adolescent (10-14 years) and late adolescent (15-19 years). Of those who had initiated sex, 1364 (4.3%) initiated sex during their early adolescent age while 6627 (21.1%) initiated sex during the late adolescent age. The explanatory variables (covariates) collected for the study include

Current Age at Survey: Metrical

The categorical covariates include:

Location: Urban or Rural (Reference)

Zone: North Central (Reference), North East, North West, South East, South South, South West.

Educational Attainment: No education (Reference), Primary, Secondary or Higher.

Religion: Traditional (Reference), Christian or Islam

Wealth Index: Poor, Middle (Reference) or Rich

Terminated Pregnancy: Yes or No (Reference)

Listen to Radio: At most once a week (Reference) or More than once a week

Watch Television: At most once a week (Reference) or More than once a week

Current Age (Grouped): 15-19 years (Reference), 20-30 years or 31-49 years

State: State in Nigeria where the respondent resides (see Figure 1).

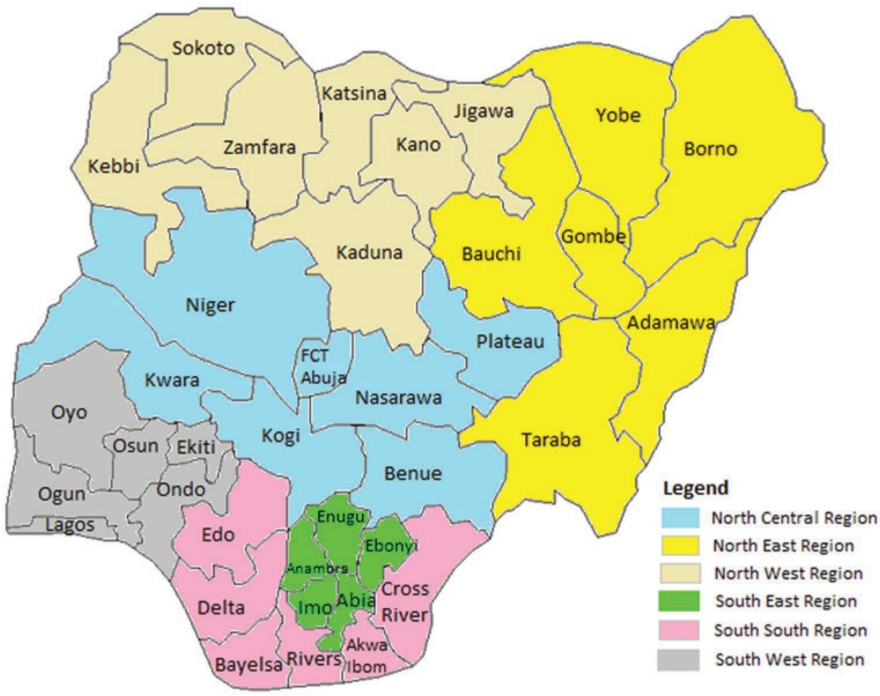

Figure 1. Map of Nigeria showing the 36 states and Federal Capital Territory (FCT), Abuja.

Binary Response Model

A binary response model applies where an event Y has two outcomes which can take on value 1 if an event of interest occurs and value 0 if the event does not occur. The distribution of Y can be specified by probabilities $\Pr(Y=1)=\pi$ and $\Pr(Y=0)=1-\pi$ with mean given as $E(Y)=\pi$ and variance $Var(Y)=\pi(1-\pi)$.

Generalized linear model (GLM), (Hastie and Tibshirani, 1990) has become one of the standard tools for analysing the impact of predictor variables on various types of response variables.

Geoadditive Binary Probit Model Framework

Let $y_1,\ldots,y_n$ denote n observable binary responses, where $y_i= 0,1$; ,i=1, . . .,n and. Geoadditive regression models extend (generalized) linear models by adding nonparametric terms for nonlinear effects of continuous covariates and geographic effects of a spatial variable to the usual linear part of the predictor.

$x_i = [x_{i1,}x_{i1},\ldots,x_{ip}]'$, denote the corresponding covariate vector for ith response. We can then model the probability of y_i by independent Bernoulli distribution which can be given as

$$y_i \sim \text{Bernoulli}(p_i), i=1,\ldots,n, \tag{1}$$

where p_i is the probability that $y_i=1$

Clearly, (1) can be connected to x_i using the probit model with the conventional linear predictors given by

$$\Pr(y_i=1 \mid x_i)=\Phi(\beta_{0i}+x_i'\beta_i), \tag{2}$$

where Φ is the standard normal distribution function.

Assumption of linear relationship between y_i and x_i as in the case of (2) is not always plausible as effects of some metrical covariates such age are not often linear on response and cannot be adequately modelled parametrically as fixed effects (Adebayo and Fahrmeir, 2005; Abiodun *et al.*, 2014). Also, small area effects may not be adequately estimated in data containing geographical information with the model. Geoadditive probit model can therefore be obtained by extending the linear predictor in (2) to a more general semi-parametric form

$$\Pr(y_i=1 \mid x_i)=\Phi(\eta_i)$$

with

$$\eta_i = \beta_{0i} + \gamma_i' x_i + f_1(z_1)\ldots f_p(z_p) + f_{spat}(s_i) \tag{3}$$

where the original $x_i'\beta_i$ in (2) is separated into f_j, the nonlinear effect of a metrical covariate z_j; γ, the vector of the usual linear fixed effects of x_i and $f_{spat}(s_i)$ the spatial effect which captures the residual variations within or between areas or locations not explained by other components of the models.

One major advantage of the geoadditive model given in (3) is that it allows the simultaneous incorporation of small area spatial effects, non-linear effects of the metric covariates as well as the usual linear fixed effects of categorical variables in a unifying model framework.

One may further split up spatial effects f_{spat} into spatially correlated (structured) and uncorrelated (unstructured) effects as

$$f_{spat}(s_i) = f_{strt}(s_i) + f_{unstr}(s_i) \tag{4}$$

A rationale behind this is that a spatial effect is a surrogate of many unobserved influential factors, some of which may be a strong spatial structure and others may only be present locally (Adebayo and Fahrmeir, 2005).

Bayesian Inference

In Bayesian context, all parameters and function evaluations are considered as random variables upon which appropriate priors are assumed. In this study, independent diffuse prior was assumed for the parameters of fixed effects γ. For the non-linear effects f_j, a Bayesian P-spline prior as in Lang and Brezger (2004) and Brezger and Lang (2006) was assumed.

Omitting the indices, each function f can be approximated by a linear combination

$$f(z) = \sum_{j=1}^{J} \beta_j B_j(z) \tag{5}$$

of B-splines basis functions and the smoothness of function f is achieved by penalizing differences of coefficients of adjacent B-splines (Eilers and

Marx, 1996) or equivalently, by assuming second order Gaussian random walk smoothness prior

$$\beta_j = 2\beta_{j-1} - \beta_{j-2} + u_j \quad (6)$$

with Gaussian errors $u_j \sim N(0,\tau^2)$, e.g Fahrmeir and Lang (2001). The variance parameter τ^2 controls the amount of smoothness and it can be estimated from the data.

For the structured spatial effects $f_{str}(s)$ in (4), the Gaussian Markov random field prior (MRF), which is common in spatial statistics has been chosen as in Besag *et al.* (1991).

These prior reflects spatial neighbourhood relationship. For data with geographical information, two regions, say s and r are neighbours if they share a common boundary. Then a spatial extension of random walk models leads to the conditional, spatially autoregressive specification

$$\{f_{str}(s) \mid f_{str}(r); r \neq s, \tau^2\} \sim N\left(\sum_{r \in \partial_s} f_{str}(p)/N_s, \tau^2/N_s\right) \quad (7)$$

where N_s is the number of adjacent sites and $sr\partial\in$ denotes that region r is a neighbour of region s. Again τ^2 controls the amount of spatial smoothness.

To estimate the smoothing parameter τ^2, a highly dispersed but proper hyper-prior was assigned to τ^2. Hence for all variance components, an inverse gamma distribution with hyper-parameters a and b was chosen, e.g. $2\tau \sim IG(a, b)$. Standard choices for the hyper-parameters are a=1 and b=0.005 or a=b=0.001.

Fully Bayesian inference was based on the posterior distribution of the model parameters, which is not of a known form. Therefore, Markov Chain Monte Carlo (MCMC) sampling from full conditionals for nonlinear effects, spatial effects, fixed effects, and smoothing parameters was used for posterior analysis. For model comparison and selection, Deviance Information Criterion (DIC) of Spiegelhalter *et al.*(2002) was used.

In analyzing the data using probit model, the response variable was indicated as follows

$$y_i = \begin{cases} 1 \text{ if a woman initiated sex at early adolescent age} \\ 0 \text{ otherwise} \end{cases} \quad (8)$$

Geoadditive probit models was fitted to the data based on the indicator in (8) as

$\Pr(y_i=1 \mid x_i) = \Phi(\eta_i)$,

with

$$\eta_i = \beta_{0i} + \gamma_i' x_i + f(age) + f_{spat}(state)$$

where

γ is a vector of fixed effects of categorical variables, $f(age)$ is a nonlinear effect of current age of women at survey and $f_{spat}(state)$ is the spatially correlated effect of state.

Model Selection

Six models were progressively considered to investigate the combination of variables that best explain adolescent age at sexual initiation. In first model (M_1), all the variables, including the metrical age were fitted linearly as fixed effect. The second model (M_2) included the spatial effect of state to M_1. In the third model (M_3), we fitted all the categorical variables were fitted linearly as fixed effect and age as nonlinear effect. In model (M_4), spatial effect of state was included in model M_3. The fifth model (M_5) contained all the variables as categorical including categorized age while model M_6 included spatial effect in model M_5. The six models were implemented in BayesX version 2.1

RESULTS

Table 1 contains the effective numbers of parameters (PD) and Deviance Information Criterion (DIC) for the six progressively fitted models. The best model based on least DIC of 6496.389 is M_4.

Table 1: The Deviance Information Criterion (DIC)

Model	Model Description	pD	DIC
M_1	Categorical variables including age metrical as fixed effect	19.602	6671. 473
M_2	Categorical variables including current age metrical as fixed effect +spatial effects of state	43.076	6526.843
M_3	Categorical variables as fixed effects+nonl inear effect of current age	26.746	6635.477
M_4	Categorical variables as fixed effects+nonlinear effect of age + spatial effect of state	50. 808	6496.389
M_5	All variables including current age categorized as fixed effect	18.140	6672.124
M_6	All variables including current age categorized as fixed effect + spatial effects of state	42.291	6527.647

Results of Fixed Effects

The regression coefficients (posterior means) for the six models are found to be similar (results not shown). However, one major interest of this study is to investigate the association between the current age groups of Nigerian women and their adolescent age at sexual initiation. Therefore the discussions of the fixed effects have been based on model M_6 that contains the age groups. The posterior estimates (means) along with the standard errors and 95% credible intervals are presented in Table 2. As observed, women who are urban dwellers are less likely to initiate sex during their early adolescent age compared to the rural dwellers. Women who ever had terminated pregnancy are more likely to initiate sexual activities during early adolescent age than those who never experienced terminated pregnancy. Also, women who are currently below 20 years of age are at higher risk of initiating sex during early adolescent age than those currently in the higher age groups 20 – 30 and 31-49 years. Women who have at least primary education have lower likelihood of initiating sex during early adolescent age compared to those who have no education education. The upper quintile category of wealth index (rich) is associated with significant lower risk of early adolescent sexual initiation while lower quintile category is associated with higher, though non-significant risk. The analysis also shows that listening to radio and watching television have little or no association with early adolescent sexual initiation.

Table 2: Posterior Estimates of the Fixed Effects Covariates of Model M6

Variable	Coefficient	Std. Error	95% C.I Lower	Upper
Constant	-0.291[a]	0.149	-0.598	-0.004
Location				
Rural (Ref)	0			
Urban	-0.195[a]	0.081	-0.356	-0.035
Terminated Pregnancy				
No (Ref)	0			
Yes	0.256[a]	0.091	0.077	0.436
Current Age (years)				
15-19 (Ref)	0			
20-30	-0.930[a]	0.118	-1.162	-0.697
31-49	-1.224[a]	0.128	-1.475	-0.972
Educational Attainment				
None (Ref)	0			
Primary	-0.408[a]	0.087	-0.580	-0.237
Secondary	-0.934[a]	0.096	-1.123	-0.746
Higher	-1.308[a]	0.105	-1.727	-0.889
Zone				
North Central(Ref)	0			
North East	0.197	0.112	-0.023	0.417
North West	0.705[a]	0.132	0.446	0.966
South East	0.148	0.141	-0.127	0.423
South South	0.564[a]	0.108	0.351	0.777
South West	-0.036	0.127	-0.287	0.214
Wealth Index				
Middle (Ref)	0			
Poor	-0.007	0.084	-0.172	0.158
Rich	-0.426[a]	0.089	-0.602	-0.252
Listen to Radio				
At most once a week (Ref)	0			
More than once a week	-0.062	0.079	-0.217	-0.092
Watching Television				
At most once a week (Ref)	0			
More than once a week	0.017	0.089	-0.155	0.191

C. I denotes Confidence Interval

[a]Indicates statistically significant effect at 5% level

Non-linear Effects

Figure 2 shows the non-linear effect of metrical current age at survey. It is clearly observed from the figure that the relationship between adolescent age at first sex and current age of the women respondent is non-linear. An approximately inverse U-shape feature is evident, indicating the inappropriateness of considering the metrical current age as having linear relationship with adolescent age at sexual initiation.

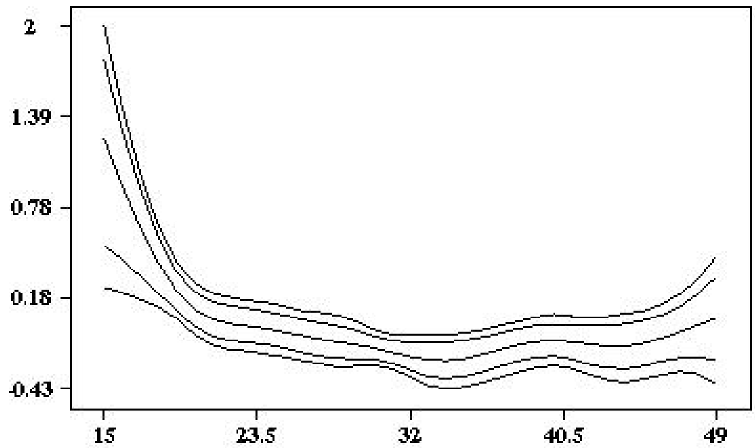

Figure 2: Non-linear effect of current age together with pointwise 80% and 95% credible intervals

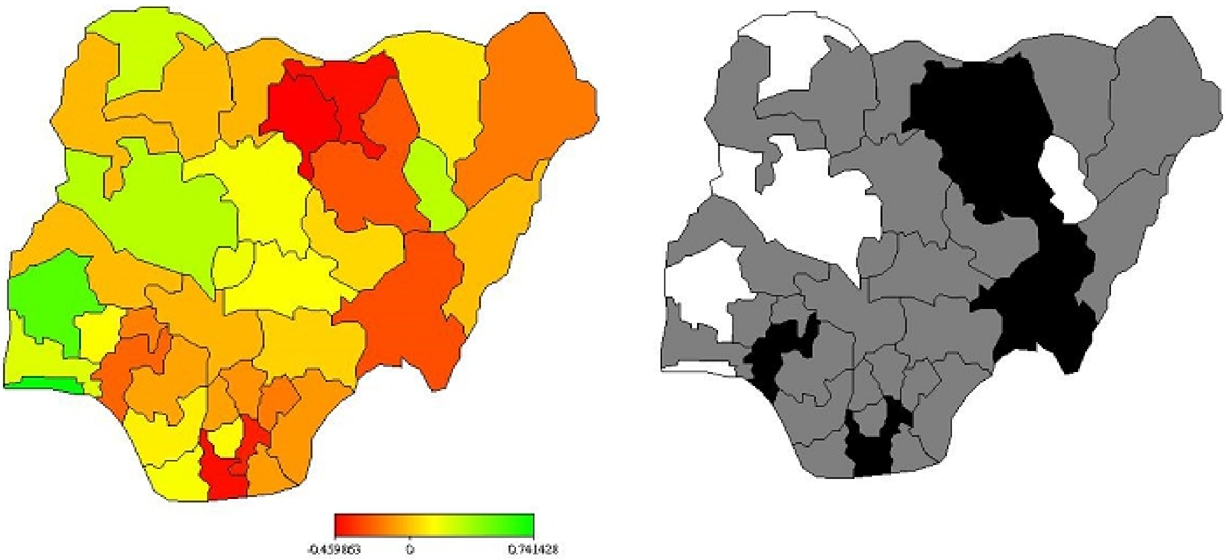

Figure 3: Spatial effects of States on Adolescent Age at Sexual Initiation Model M6.

Spatial Effects

Figure 1 displays the map of Nigeria showing the 36 states, while Figure 3 shows the posterior means with the 95% posterior probabilities of spatial effects. It is obvious that there exists a considerable amount of residual geographical variability at state level on the sexual behaviours of the Nigerian women during their adolescent age, even after controlling for the other covariates. States with black colours are associated with negatively significant spatial effects implying that the risk of early adolescent (9-14 years) sexual initiation are more in those states, while states with white colours indicate that more women who had adolescent sexual initiation had it during their late adolescent age. However, the spatial effects in states with grey colour are not significant. As observed, Jigawa, Bauchi, Taraba, River and Ondo states are associated with early adolescent sexual initiation. Late adolescent sexual initiation are experienced more by women in Niger, Sokoto, Oyo and, Gombe states, while adolescent sexual initiation is not significant in the remaining states.

CONCLUSION

Geoadditive probit model offers a novel approach of analyzing adolescent age at sexual initiation in Nigeria within a joint modelling framework. The approach allows flexibly modelling of metric covariate that has nonlinear effects as well as incorporating spatial effect in a unifying model framework. The spatial analysis clearly revealed the data structure which are often overlooked in analyses with standard regression models. The approach has also provided us an insight into the functional pattern of the respondent's current age in which case, considering age as a linear effect would have resulted in spurious and unreliable conclusion.

REFERENCES

Abiodun, A.A., Adebayo, S.B., Oyejola, B.A., Anyanti, J. and O.Ladipo. 2014. A Spatial Analysis of Age at Sexual Initiation Among Nigerian Youth as a Tool for HIV Prevention: A Bayesian Approach. Springer Dordrecht Heidelberg New York London.

Adebayo, S.B. and Fahrmeir, L. 2005. Analyzing Child Mortality in Nigeria with Geoadditive Discrete-time Survival Models. *Statistics in Medicine,* 24:709-728.

Agha, S., van Rossem, R., and Ankomah, A. 2006. Community Level Influences on Primary Abstinence in Nigeria. *DHS Working Paper.*

Andersson-Ellstrom, A., Forssman, L., and Milsom ,I. 1996. Age of sexual debut related to life-style and reproductive health factors in a group of Swedish teenage girls. *Acta. Obstet. Gynecol. Scand.* 75(5): 484-489.

BayesX- Software for Bayesian Inference based on Markov Chain Monte Carlo. simulation Techniques,Version 2.1. Available under http://www.stat.uni-muenchen.de/~bayesx

Besag, J., York, Y. and Mollie, A. 1991. Bayesian Image Restoration with two Applications in Spatial Statistics (with discussion). *Annals of the Institute of Statistical Mathematics*, 43, 1-59.

Brezger, A. and Lang, S. 2006. *Generalized Structured Additive Regre*ssion based on Bayesian P-splines. *Computational Statistics and Data Analysis* 50, 967-991.

Brewster, K. L. 1994. 'Neighborhood context and the transition to sexual activity among young Black women', *Demography* 31, 603-614.

Chiao,C. and Mishra,V. 2007. Trends in Primary and Secondary Abstinence among Kenyan youth. *DHS Working Paper* 36.

Dickson, N., Paul, C., Herbison, P. and Silva, P. 1998. First Sexual Intercourse: Age, Coercion and later Regrets reported by a Birth Cohort. *British Medical Journal*, 316, 29-33.

Ekundayo, O.J, Dodson-Stallworth, J., Roofe, M., Aban, I.B., Bachmann, L.H. and Ehiri, J. 2007.The determinants of sexual intercourse before age 16 years among rural Jamaican adolescents. *The Scientific World Journal:* Child Health and Human Development. 7(94): 493-503.

Eilers, P.H.C. and Marx, D.B. 1996. Flexible Smoothing with B-Splines and Penalties. *Statistical Science*, 11(2), 89-121.

Fahrmeir, L. and Lang, S. 2001. Bayesian Semiparametric Regression Analysis of Multicategorical Time-Space Data. *Annals of the Institute of Statistical Mathematics*, 52(1), 1-18.

Federal Ministry of Health [FMOH]. 2006. Technical Report on 2005 National HIV/Syphilis Sero-Prevalence Sentinel Survey among Pregnant Women Attending Antenatal Clinics in Nigeria. *Federal Ministry of Health, Abuja, Nigeria.*

Hastie, T. and Tibshirani, R. 1990. Generalized Additive Models. *Chapman and Hall London.*

Harrison A., Cleland J., Gouws E., and Frohlich J. 2005. Early sexual debut among young men in rural South Africa: heightened vulnerability to sexual risk? *Sex Transm Infect.* 81:259–61

Kinsman, S.B., Roma, D., Furstenberg, F.F, and Schwarz, D.F. 1998. Early sexual initiation: the role of peer norms. *Paediatrics,* 102(5), 1185-1192.

Lang, S. and Brezger, A. 2004. Bayesian P-splines. *Journal of Computational and Graphical Statistics*, 13: 183-212.

Masatu, M.C., Kazaura, M.R., Ndeki, S. and Mwampambe, R. 2009. Predictors of risky sexual behavior among adolescents in Tanzania. *AIDS and Behaviours.* 13: 94-99.

Mmbaga, E.J., Leonard, F. and Leyna, G.H. 2012. Incidence and predictors of adolescent's early sexual debut after three decades of HIV interventions in Tanzania: A time to debut analysis. *PLoS ONE 2012,*, 7(7): 1-9.

Nnebue, C. C., Chimah, U. C., Duru, B.C., L Ilika, A.L and Lawoyin, T.O. 2016. Determinants of Age at Sexual Initiation among Nigerian Adolescents: A Study of Secondary Schools Students in a Military Barracks in Nigeria. *American Journal of Medical Sciences and Medicine.* 4(1): 1-7.

http://pubs.sciepub.com/ajmsm/4/1/1

Pettifor A.E., Rees H.V., and Kleinschmidt I. 2005. Young people's sexual health in South Africa: HIV prevalence and sexual behaviors from a nationally representative household survey. 19:1525–34

Rink , E., Tricker, R., and Harvey, S.M. 2007. Onset of sexual intercourse among female adolescents: the influence of perceptions, depression and ecological factors. *J. Adolesc. Health*. 41(4), 398-406 .

Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and van der Linde, A. 2002. Bayesian Measures of Model Complexity and Fit (with discussion). *Journal of the Royal Statistical Society, B*, 64, 583-640.

The World Health Report. 2004.