INTRA-CLASS CORRELATION ON AGE VARIATION AT SEXUAL DEBUT IN NIGERIA

Supported by tetfund TERTIARY EDUCATION TRUST FUND

Adesupo A. Akinrefon*, Blessing Okoye and Olateju A Bamigbala
Department of Statistics and Operations Research, Modibbo Adama University of Technology, Yola, Nigeria
*Correspondent author: akinrefon.adesupo@mautech.edu.ng

Received: November 18, 2018 **Accepted:** March 09, 2019

Abstract: The age of initiation of sexual intercourse is an increasingly important issue to study given that sexually active young women are at risk of multiple outcomes. Knowledge about the factors associated with age at first sex plays a major role in controlling population growth rate. The purpose of this study is to examine geographical disparities in age at first sexual intercourse, and to examine individual and state – level predictors of early sexual initiation. We adopted Multilevel Linear Modeling, known for its ability to explain the effects of the shared variances present in the study on the variable of interest using survey data from 2008 and 2013 Nigeria Demographic Health Survey (NDHS), collected via a multistage clustered sampling scheme. We investigated the factors that were thought to be associated with variation in age at initiation of sex among Nigerian women. It was observed that the average age at which a Nigerian woman becomes sexually active is 17 years which is a teenage year. 35% of the variation in ages at first sex can be attributed to the geographical location (geopolitical zones) where they are resident, and 3.31% to the differences in years of the surveys, respectively.

Keywords: Sexual debut, demographic health, multilevel linear model, intra-class correlation

Introduction

The age at first sexual intercourse is an increasingly important subject to study owing to the fact that sexually active young women are usually at risk of multiple outcomes including early pregnancies, sexually transmitted infections and Vesico-Vaginal fistula (VVF). To date, most studies on the effects of early initiation of sexual intercourse among women have focused on those who had their first sexual experience before marriage. The age of a person's first sexual experience can influence their future development and relationships. Previous studies (Orji and Esimai, 2005; Dongurum *et al*., 2009) have established the prevalence of sexual activities among adolescence less than 15 years of age. The inability to make informed decision regarding family planning and to take adequate protective measures during coitus predisposes this young people to risks of infections. Many researchers have focused on the demographic, family, and social factors associated with sexual initiation and reasons adolescents begin having consensual intercourse. Less is known, however, about the geographical and contextual factors associated with age of first sexual intercourse. The analysis of geopolitical variation in age of first sexual experience is important.

Methodology

Nigeria is located in Western Africa on the Gulf of Guinea and has a total area of 923,768 kilometre squared (km^2). It shares land border with the Republic of Benin in the west, Niger in the North and Chad and Cameroun in the East. Nigeria is often referred to as the "Giant of Africa" because it is the most populous country in Africa. The United Nations estimates that the population of Nigeria is currently 186,485,918 as of Friday, May 27, 2016, based on the latest United Nations estimates, with 48.1% of the population distributed as urban (91,668,667 people in 2016) and 51.9% rural and The population density in Nigeria is 205 per Km^2 (532 people per m^2). Nigeria's population is equivalent to 2.48% of the total world population (United Nations, 2016).

Owing to the hierarchical nature of demographic and health data, the survey design that was used is the multistage cluster sampling design using data from the 2008 and 2013 Nigeria Demographic and Health survey (NDHS). The sample was designed to provide population and health indicator estimates at the national, zonal, and state levels.

Multilevel multiple regression models were applied to data on 57226 ever or currently married women who had participated in 2008 and 2013 Nigeria Demographic and Health Survey. The age at first sexual intercourse was defined as the dependent variable thus, if a woman has never had sexual intercourse in her life (e.g., a "virgin") and never married, she was not included in the analysis. Factors affecting Variation in age at first sexual intercourse was also studied. Four variables indicating individual-level features thought to affect age of initiation at first sexual intercourse were included in this study. The variables are type of residence, highest educational level, wealth index and religion affiliation. Also geopolitical zone was included on the second level while year of survey was introduced in the third level.

The multilevel regression model

The multilevel regression model also known as mixed linear model, random coefficient model, hierarchical linear model or variance component model in the Literature assumes hierarchical data with the response variable measured at lowest level while the explanatory variables can exist at all levels.

Estimating the parameters

In a two-level multilevel regression model with levels 1 and 2, the dependent variable y_{ij} and the explanatory variables x_{ij} are measured at the lowest level while the explanatory variable $z_j's$ are measured at the second (higher) level. Separate level 1 models are developed for each level 2 units. These models are also called within-units models as they describe the effect in the context of a single group (Gill, 2003). The separate regression equation for each group is given as

$$y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + e_{ij} \qquad (1)$$

Where y_{ij} is the dependent variable measured for the ith level 1 unit nested within the j^{th} level 2 unit, x_{ij} is the value on the level 1 prediction, β_{0j} is the intercept for the j^{th} level 2 unit, β_{1j} is the regression coefficient associated with x_{ij} for the j^{th} level 2 unit and e_{ij} is the random error associated with the i^{th} level 1 unit nested within the j^{th} level 2 unit.

The regression coefficients carry a subscript j indicating that may vary across the level 2. These are modelled by explanatory variables and random residual term at the level 2.

Level 2 models are also referred to as between-unit models as they describe the variability across multiple levels (Gill, 2003). Consider the case of a single level 2 predictor, the model is given as

$$\beta_{0j} = \gamma_{00} + \gamma_{01} z_j + u_{0j} \quad (2)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} z_j + u_{1j} \quad (3)$$

Where z_j is value on the level-2 predictor, γ_{00} =is the overall mean intercept adjusted for z, γ_{10} is the overall mean intercept adjusted for z, γ_{01} is the regression coefficient associated with z relative to level-1 intercept, γ_{11} is the regression coefficient associated with z relative to level-1 slope, U_{0j} is the random effects of the j^{th} level-2 unit adjusted for z on the intercept and U_{1j} is the random effects of the j^{th} level-2 unit adjusted for z on the slope.

Substituting equations (2) and (3) into equation 1, it gives a single-equation of the multilevel regression model

$$y_{ij} = \gamma_{00} + \gamma_{10} x_{ij} + \gamma_{01} z_j + \gamma_{11} z_j x_{ij} + u_{ij} x_{ij} + u_{0j} + e_{ij} \quad (4)$$

If there are P-explanatory variables at the lowest level and Q-explanatory variables at the higher level, then, equation (4) become

$$y_{ij} = \gamma_{00} + \sum_p \gamma_{p0} x_{pij} + \sum_q \gamma_{0q} z_{qj} + \sum_q \sum_p \gamma_{pq} z_{qj} x_{pij} + \sum_p u_{pj} x_{pij} + u_{0j} + e_{ij} \quad (5)$$

The γ are the regression coefficients, u are the residuals at the group level and the e is the residual at the lowest level. The above can be generalized for more than 2 levels.

The proportion of variance in the population explained by the clustering structure is given by the intra-class correlation ρ which is estimated using the null model, that is, the model with no explanatory variable, called the intercept-only model;

$$y_{ij} = \gamma_{00} + u_{0j} + e_{ij} \quad (6)$$

The intra-class correlation ρ is estimated by the equation:

$$\rho = \frac{\sigma^2_{u0}}{\sigma^2_{u0} + \sigma^2_e + \sigma^2_{v0}} \quad (7)$$

where σ^2_e is the variation at lowest level, σ^2_{uo} is the variation at the second level and σ^2_{vo} is the variation at the third level.

The analysis was done in three steps. In Model 1 (null model); In Model 2, only individual-level factors were included, while in Model 3, state-level factors were added to Model 2. The results of fixed effects (measures of association) and random effects were reported.

Restricted maximum likelihood (REML)

REML includes only the variance components, that is, the parameter that parameterize the random-effect terms in the linear mixed-effect model.

The REML methods compare two models that are nested in their random effect terms with the same fixed effects design. RML yields better estimates of the variance component than the Full Maximum Likelihood.

A multilevel model of the form

$$y = \underset{fixed}{x\beta} + \underset{random}{zb} + \underset{error}{\varepsilon}, \quad (8)$$

Where y is the n by 1 response vector and n is the number of observation, x is an n by p fixed effects design matrix, β is a p by 1 fixed effects vector, z is an n by q random design matrix, b is a q by 1 random effect vector and ε is the n by 1 observation error vector. The random-effects vector, b, and the error vector, ε, are assumed to the prior distributions;

$$b \sim N(0,\ \sigma^2 D(\theta)), \qquad \varepsilon \sim N(0,\ \sigma^2 I),$$

Where D is the symmetric and positive semi-definite matrix, parameterized by a variance component vector θ, I is an n *by* n identity matrix, and σ^2 is the error variance.

In this model, the parameters to estimate are the fixed-effects coefficients β, and the variance components θ and ε.

β is estimated in a second step. Assuming a uniform improper prior distribution for β and integrating the likelihood $L(y \mid \theta, \sigma^2)$ with respect to β results in the restricted likelihood $L(y \mid \theta, \sigma^2)$.

$$L(y \mid \theta, \sigma^2) = \int L(y \mid \beta, \theta, \sigma^2) L(\beta) d\beta = \int L(y \mid \beta, \theta, \sigma^2) d\beta$$

The algorithm first profiles out $\hat{\sigma}^2_R$ and maximizes the remaining objective function with respect to θ to find $\hat{\theta}_R$. The restricted likelihood is then maximized with respect to σ^2 to find $\hat{\sigma}^2_R$. Then, it maximizes β by finding its expected value with respect to the posterior distribution.

Results and Discussion

Table 1: Frequency and percentage distribution of background characteristics

Characteristics	2008		2013	
	Frequency	%	Frequency	%
Geopolitical Zones				
North Central	4621	17.6	4566	14.8
North East	5260	20.0	5486	17.7
North West	6263	23.8	8276	26.8
South East	2456	9.3	3202	10.4
South South	3845	14.6	4775	15.4
South West	3868	14.7	4608	14.9
Highest Educational Level				
No Education	11697	44.5	12306	39.8
Primary Education	5328	20.2	5943	19.2
Secondary Education	7146	27.2	9577	31.0
Higher	2142	8.1	3087	10.0
Religion				
Catholic	2558	9.7	3046	9.9
Other Christian	10211	38.8	12088	39.1
Islam	13094	49.8	15499	50.1
Traditionalist	450	1.7	280	0.9
Type of Residence				
Urban	7821	29.7	11641	37.7
Rural	18492	70.3	19272	62.3
Wealth Index				
Poorest	6267	23.8	5790	18.7
Poorer	5558	21.1	6209	20.1
Middle	5046	19.2	6177	19.9
Richer	4940	18.8	6384	20.7
Richest	4502	17.1	6353	20.6

Table 2 presents the prevalence of early sex initiation in Nigeria. It presents the intercept only model to determine the average age at which Nigerian women first indulge in sexual intercourse.

As can be observed from the table 2 below, the average age at which a Nigerian woman becomes sexual active is approximately 17 years.

The intercept only model is given as;

$$(Age\ at\ First\ Sex)_{ijk} = y_{000} \quad (9)$$

Where y_{000} is the grand mean.

Table 2: Intercept only model

Age at First Sex	Coefficient	Std. Error	Z	P>|Z|
cons	17.13573	0.1835733	93.3454	0.000

Table 2 presents the Null Model with random intercept. It provides some useful information needed to calculate the intra-class correlation coefficient between the states and the year, which will help us understand the structure of the data.

The intra-class correlation coefficient represents the percentage variation in age at first sex that is attributive to the state where the woman is domicile, and year of survey. The standard deviation of each individual on age at first sex from its state mean is estimated as 3.173744 while the standard deviation of each state mean within its zone is 1.574614; also, the standard deviation of each year from its grand mean is 0.1524139 and the overall grand mean is 17.13573 as seen in Tables 3a&b.

The method illustrated by Siddiqui *et al.* (1996) which defines the intra-class correlations at the zone and survey-year level, was used to calculate the intra-class correlation coefficient (ICC), as:

$$\rho_{state} = \frac{\sigma_{v_0}^2 + \sigma_{u_0}^2}{\sigma_{v_0}^2 + \sigma_{u_0}^2 + \sigma_e^2} \quad (10)$$

And

$$\rho_{year} = \frac{\sigma_{v_0}^2}{\sigma_{v_0}^2 + \sigma_{u_0}^2 + \sigma_e^2} \quad (11)$$

There is significant variation in the age of early sex debut across the zones ($\sigma_{u_0}^2$= 1.574614, p – value= .000) and survey-year ($\sigma_{v_0}^2$ = 0.1524139, p – value = .000). Using (10) and (11), the intra – class correlation at zone and survey-year level are estimated as 0.3524 and 0.0312, respectively. This intra-Class correlation coefficient results implied that about 3.13% of the variation in the age of first sex can be attributed the difference in year survey was conducted, and about 35% of the variation in the age of initiation of sexual intercourse arises from the differences in the state-level factor (which in this case is the geo – political zones). The 3.13% component also indicates that there is a 3.13 percent change in the age at first sex between the survey of 2008 and 2013. This variation remained significant, even after controlling for individual-level factors (Model 3).

Table 3a: Intercept only model with random effect

Age at first sex	Coefficient	Std. Error	Z	P>|Z|
_cons	17.13573	0.183573	93.345	0.000

Table 3b: Intercept only model random-effects parameters

Random-effects parameters	Estimate	Std. Error
Year: Standard deviation	0.1524139	3.55053
State: Standard deviation	1.565735	3.187224
Individual: Standard deviation	3.7370	0.0313

Modeling age at first sex

The age at first sex was modeled using multilevel multiple regression and the results are presented in the Table 4 below. The table contains various models generated and their AIC's. Table 4 shows a decline in the AIC value of the different models. The decline in AIC value indicates that the inclusion of factors at the second and third level contributed significantly to the model. Model 1 (the null (fixed) model, (intercept only model)) while in model 2, a random intercept model was considered. In models 3 and 4, the various predictor variables were considered.

Table 4: The multilevel multiple regression models for age at first sex with their AIC's

Model	Models Generated	Equations	AIC's
1.	Null Model	$y_{ijk} = \gamma_{00} + \mu_{0j} + \varepsilon_{ijk}$	295291
2.	Random Intercept Model	$y_{ijk} = \gamma_{000} + \gamma\mu_{xij} + \mu_{0j} + \varepsilon_{ijk}$	294977
3	Model with individual factors.	$y_{ijk} = \alpha + b_1 Typeofres + b_2 Zones + b_3 Wealthindx + b_4 HighestEdulv + b_5 Religion$	291232
4.	Individual + 2nd Level factors	$y_{ijk} = \alpha + b_1 GeopolZone + b_2 HighestEduLv + b_3 Typeofres + b_4 Religion + b_5 Wealthindx$	289501

Table 5 presents the random intercept model with the individual level predictors, no variable was declared for the state and the year levels. From the result, it was observed that for a change in type of residence, i.e., if persons are from urban centers, the age at first sex decreases by 0.203 if other factors are held constant. The negative sign indicates that women in urban settlements delay in age at sexual initiation compared to rural dwellers. A woman from a very rich background has a 0.979 increase in age at first sex, a woman of average wealth index has a 0.278 increase in age at first sex, and a woman from a poor background age at first sex is increased by 0.176, if all other factors are held constant. All these goes to indicate that the chance of an earlier sexual debut seems to decrease with increase in wealth index implying that poverty seem to be a major factor that prompt early sexual debut. For women of Catholic background, the chance of early sexual engagement is 0.546, women of other Christian religion background have 0.511, while traditional religion women are at 0.103 where Muslim women are used as reference. This implies that of all the different religion, women of catholic background have some chance of a much later age of sexual debut followed by other Christian faith, while Islamic women have the earliest. The chance of early first sex for women who have higher education is 2.646, that of women with secondary education is increased by 0.714. If a woman has no education, her age at first sex is decreased by 0.276, which give her a chance of an earlier sex debut. These are in comparison to women with only primary education. Uneducated women seem to engage more in early sexual debut than others.

Table 6 presents the varying intercept and varying coefficient model with the inclusion of the state level (level two) predictor (Geopolitical Zone), the inclusion of geopolitical zone at the state level reported the variable, type of residence to be significant which increase the age at first sex of rural women by 0.201 when compared to urban women as against the 0.204 when the state level predictor was not included while the marginal effect of being a traditionalist still remained insignificant to age at first sex compared to being an Islam.

Table 5a: Varying intercept model with individual level predictors

Parameter	Age at first sex	Coefficient	Std. Error	P>|Z|
Type of Residence	Rural	-0.203	0.036	0.000*
Wealth Index	Poorer	0.176	0.0407	0.000*
	Middle	0.278	0.046	0.000*
	Richer	0.485	0.052	0.000*
	Richest	0.979	0.062	0.000*
Religion	Catholic	0.546	0.062	0.000*
	Traditionalist	0.103	0.112	0.382
	Others	0.511	0.042	0.000*
Highest	No education	-0.276	0.042	0.000*
Educational	Secondary	0.714	0.039	0.000*
Level	Higher	2.646	0.057	0.000*
	Constant	15.195	0.175	0.000*

Note: Asterisk (*) indicate the variables that are significant at 5% level of significance

Table 5b: Parameters of varying intercept model with individual level predictors

Random-effects parameter	Estimate	Std. error
year: Standard deviation	0.017	0.120
State: Standard deviation	0.975	3.027
Individual: Standard deviation	2.606	0.129

Table 6a: Varying intercept and varying coefficient model with the inclusion of the state level (level two) predictor (geopolitical zone)

Parameter	Age at first sex	Coefficient	Std. Error	P>|Z|
Type of Residence	Rural	-0.201	0.036	0.000*
Wealth Index	Poorer	0.173	0.0406	0.000*
	Middle	0.270	0.046	0.000*
	Richer	0.479	0.052	0.000*
	Richest	0.974	0.062	0.000*
Religion	Catholic	0.511	0.062	0.000*
	Traditionalist	0.085	0.117	0.4713
	Others	0.489	0.044	0.000*
Highest	No education	-0.270	0.042	0.000*
Educational	Secondary	0.714	0.039	0.000*
Level	Higher	2.648	0.057	0.000*
	Constant	15.144	0.197	0.000*

Note: Asterisk (*) indicate the variables that are significant at 5% level of significance

Table 6b: Parameters of varying Intercept model with individual level predictors

Random-effects Parameter	Estimate	Std. Error
year: Standard deviation	0.017	0.120
State: Standard deviation	0.614	3.027
Individual: Standard deviation	2.580	0.175

Summary

A study was carried out on the age at first sex of Nigerian women using dataset from 2008 and 2013 Nigeria Demographic and Health Survey (NDHS). For the study, a three-level model which account for hierarchical structure of the data was used. 35 per cent of the variation in age at first sex of Nigeria women results from differences among the states and 3.31 per cent from differences among the zones. By the explicit multilevel analytic data analysis framework, the study has shown that both individual structure and state characteristics are important predictors of women's age at first sexual intercourse, and demonstrates geopolitical variation in rates of early initiation of sex in Nigeria. After adjusting for both individual-level, state - level and year factors, the probability of starting sex at an earlier age was associated with respondents' residence type, educational level, wealth index, religion and geopolitical zone. This research has offered an alternative to more traditional ways of thinking about the distribution of sexual behaviour and sexual risk at the population level.

The potential of multilevel modeling for investigating population relative influences on individual sexual behaviour is clearly established. Nigeria is made up of six major geopolitical regions. It is ethnically and religiously diverse and economic development and education levels vary widely across the country. Not unexpectedly, this study found that Nigerian women's age of initiation of sexual intercourse varies widely by state. Specifically, traditional and cultural beliefs that seem to encourage early marriage and lead to early initiation of sexual intercourse may be principal reasons for the observed geographical disparities. The study found that women living in neighborhoods with low standard of living (rural areas) had higher chances of reporting early initiation of sexual intercourse than those in neighborhoods with higher standard of living (urban areas). This is consistent with the preponderance of evidence suggesting that Northern regions of Nigeria have some of the highest rates of early marriage in the world.

Conflict of Interest

Authors declare no conflict of interest

References

Abalkhail BA 1995. Adolescent pregnancy: Are there biological barriers for pregnancy outcomes? *The J. Egyptian Public Health Assoc.* 70(5–6): 609–625.

Adeniyi OI & Akinrefon AA 2014 Variation in Age at First Marriage among Women in Nigeria: A Multilevel Logistic Analysis. *The J. Math. Theory and Modeling*, 4(3): 121 – 127.

Adetoro OO, Babarinsa AB & Sotiloye OS 1991. Socio-cultural factors in Adolescent Septic Illicit Abortion in Ilorin, Nigeria. *Afr. J. Med. and Med. Sci.*, 20(2): 1-53.

Ethington CA 2000. Influences of the normative environment of peer groups on community college students' perceptions of growth and development. *Res. in Higher Edu.*, 41: 703-722.

Fatusi AO & Blum RW 2008. Predictors of early sexual initiation among a nationally representative sample of Nigerian adolescents. *BMC Public Health* 8: 136.

Forste RT & Heaton TB 1988. Initiation of sexual activity among female adolescents. *Youth and Society*, 19(3): 250-68.

NPC 2009. Nigeria Demographic and Health Survey 2008. Calverton, Maryland: National Population Commission and ORC/Macro.

NPC 2008. Nigeria Demographic and Health Survey. Calverton, Maryland: National Population Commission and ORC/Macro.

Otoide VO, Oronsaye F & Okonofua FE 2001. The Nigerian adolescent seek abortion rather than contraception: Evidence form focus group discussions. *Int. Family Planning Perspective*, 27(2): 77-81.

Oyediran KA, Ishola GP & Adewuyi AA 2002. Knowledge of possible pregnancy at first coitus: A study of in-school adolescents in Ibadan, Nigeria. *J. Biosocial Sci.*, 34(2): 233-48.

Oyefara IJL 2009. Socio-Economic Consequences of Adolescent Childbearing in Osun State, Nigeria.