Ilorin Journal of Science
Volume 4, Number 2, 2017, pp. 215 – 225 (Printed in Nigeria)
ISSN: 2408 – 4840 © 2017 Faculty of Physical Sciences, University of Ilorin
https://doi.org/10.54908/iljs.2017.04.02.002

ILORIN
JOURNAL OF SCEINCE

ILJS-17-002

Determinants of Desired Number of Children per Woman in Nigeria

Adeniyi[1*], O. I. and Akinrefon[2], A. A.
[1]Department of Statistics, University of Ilorin, Ilorin, Nigeria.
[2]Department of Statistics and Operations Research, Modibbo Adama University of Technology, Yola, Nigeria.

Abstract

Knowledge of the factors associated with increase in the population of country plays a key role in curtailing it leading to its control. This article presented a three-level regression analysis, using survey data from 2013 Nigeria Demographic Health Survey (NDHS). It investigated individual factors that were thought to be associated with the desired number of children by women who have not given birth and also assessed the effect of the hierarchical structure of the data. The model provided parameter and variance estimates at all levels. It was evident that variation in the desired number of children stems out from the hierarchical structure of the data. The varying intercept with individual level predictor model was better in predicting the variation in the desired number of children than varying both intercept and coefficient model. Findings show that women with lower status in terms of education and wealth index have higher desire to have more children than women with better status.

Keywords: Desired Number of Children, Multilevel Regression, Intra-class Correlation, Hierarchical Data, Akaike Information Criterion (AIC)

1. Introduction

Nigeria, most populous Africa Country accounts for approximately one sixth of the African population with 50 percent of her dwellers in the urban areas, being one of the fastest growing countries in the world with an estimated population of 140 million with an annual population growth rate of 2.9% (NPC 2006) is faced with matching population growth with development towards improved welfare, human development and economic growth. Rapid population growth in Nigeria is equally associated with unemployment with figures ranging from 17 percent per annum for the entire population to 60 percent for the youths. The reason being that, job opportunities are fewer than the number seeking for them, and stagnating economic performance. A large proportion of available resources is being consumed instead of being invested to generate growth (Federal Republic of Nigeria, 2004).

*Corresponding Author: Adeniyi, O. I.
Email: niyikiitan@yahoo.com

While it is accepted that population can be an asset for development, the truth is that Nigeria has enough mouths to fill already and there is an urgent necessity to curtail our high population growth rate. The number of children desired by each individual should be curtailed to check the booming population of the country for sustainability and developmental growth of the nation.

The desire to have children by the citizenry of a country affects the population size, particularly in Africa tradition and religion belief, where there is tendency for some women to attribute the number of children they want to have to God (Mbiti, 1970; Lee and Miller, 1990). The Yorubas in Nigeria regard children as God's gifts or blessing from heaven that cannot be refused (Olusanya, 1971). Desired family size has defined by various people is in different form, (Thompson, 2001) defined it as the number of children wanted in one's lifetime and can be viewed as the demand for children. Desired family size is the number of children parents would have if there were no subjective or economic problem involved in regulating fertility (McClelland, 1983; Brown, 2011). In order to have a nation with effective planning that can cater for the needs of her inhabitants, individuals must be ready to tailor their desire to have children to what will enhance a sustainable environment for all. National resources for development are not infinite and it is not possible to sustain high standard of living in a densely populated country like Nigeria. To achieve this in line with the goal of the National Policy on Population for Sustainable Development aimed at reduction in the total fertility rate by at least 0.6 children every five years (National Population Commission, 2004). Younger generation (that is, women who have not started giving births to children) should be educated and get aware of the menace of having uncontrollable number of children without considering the state of the economy of the country. To this end, the article focuses on the women who are yet to have children in the NDHS (2013) data.

Studies on the booming population of the developing countries have attracted great attention in recent times. To do this, the population structure of the country must be put into consideration, ignoring the population structure could possibly lead to obtaining a biased estimate of the standard error and hence the results will be misleading. Hierarchical Linear Modeling (HLM) is a complex form of ordinary least squares (OLS) regression that is used to analyse variance in the outcome variables when the predictor variables are at varying hierarchical levels. HLM accounts for the shared variance in hierarchically structured data. It accurately estimates lower-level slopes and their implementation in estimating higher-level outcomes (Hofmann, 1997). HLM simultaneously investigates relationships within and between hierarchical levels of grouped data, thereby making it more efficient at accounting for variance among variables at different levels than other existing analyses (Heather *et al*., 2012). Multilevel model is known by several names, such as 'hierarchical linear model' (Raudenbush and Bryk, 1986; Raudenbush and Bryk, 2002), 'variance component model' (Longford, 1987), and 'random coefficient model' (De Leeuw and Kreft, 1986; Longford, 1993). HLM can be ideally suited for the analysis of nested data because it identifies the relationship between predictor and outcome variables, by taking all level regression relationships into consideration.

2. The Multilevel Regression Model

Hierarchical linear model also known as 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.

2.1 Estimation of the Parameters

This non-technical description of the estimations procedures for multilevel models is largely based on Hox (2010). Multilevel models are normally estimated by Maximum Likelihood (ML), Restricted Maximum Likelihood (RML) or Iterative Generalized Least Squares (IGLS) algorithms. In the full information, using ML method, both the regression coefficients and the variance components are included in the likelihood functions. In the RML method, only the variance components are included in the likelihood function, and the regression coefficients are estimated in a second step. The RML method seems to produce less biased estimates of the variance components, especially in small samples. The difference between the two estimation methods is normally insignificant. The ML method is still used because it has some other advantages over the RML method. It is computationally easier and, since the regression coefficients are included in the likelihood functions, likelihood ratio tests can be used to compare nested models that differ in the fixed part, i.e. the number of regression coefficients.

A two-level multilevel regression model, the response (outcome) 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. The multilevel regression at the lowest level;

$$y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + e_{ij}, \tag{1}$$

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

The regression coefficients carry a subscript j indicating that they 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). Predicting with a single level 2 predictor, the model is given as

$$\beta_{0j} = \gamma_{00} + \gamma_{01} z_j + u_{0j}, \tag{2}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} z_j + u_{1j}, \tag{3}$$

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 jth level-2 unit adjusted for z on the intercept and U_{1j} is the random effects of the jth level-2 unit adjusted for z on the slope. The regression coefficient γ do not vary across groups and the between group variation left on the β coefficient assumed to be residual error is captured by the residual error term $u_{.j}$ Substituting equations 2 and 3 into equation 1, it gives a single-equation of the multilevel regression model

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

If there are p-explanatory variables at the lowest level and q-explanatory variables at the higher level, then, equation (4) becomes

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

The *γ's* are the regression coefficients, *u's* are the residuals at the group level and the e is the residual at the lowest level. A correlation showing the proportion of the variance in the response variable that stems from the variation between the higher level units is given by the Intra-class correlation coefficient (ICC) denoted as ρ . The ICC evaluates whether or not the higher level variation is ignorable or not.

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

$$\rho = \frac{\sigma^2{}_{u0}}{\sigma^2{}_{u0} + \sigma^2{}_e}, \tag{6}$$

$$u_{0j} \sim N\left(0, \sigma_{u0}{}^2\right) \quad , \qquad e_{ij} \sim N\left(0, \sigma_e{}^2\right),$$

where $\sigma^2{}_e$ is the variation at lowest level and $\sigma^2{}_{u0}$ is the variation at the second level. A multilevel model is of the form:

$$y = \underbrace{x\beta}_{fixed} + \underbrace{zb}_{random} + \underbrace{\varepsilon}_{error}, \tag{7}$$

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

$$b \sim N\left(0, \sigma^2 D(\theta)\right), \qquad \varepsilon \sim \mathrm{N}\left(0, \sigma^2 I\right),$$

where D is a symmetric and positive semi-definite matrix, parameterized by a variance component vector θ, *I* is an *n x 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*.

2.2 Restricted Maximum Likelihood (REML)

REML includes only the variance components; that is, the parameter that parameterise the random-effect terms in the linear mixed-effect model. β is estimated in a second step. Assuming a uniform improper prior distribution for β and integrating the likelihood $L(y|\theta,\sigma^2)$ with respect to β results in the restricted likelihood $L(y|\theta,\sigma^2)$.

$$L(y \mid \theta, \sigma^2) = \int L(y \mid \beta, \theta, \sigma^2) L(\beta) d\beta = \int L(y \mid \beta, \theta, \sigma^2) d\beta . \quad (8)$$

The algorithm first profiles out $\hat{\sigma}^2{}_R$ and then maximize 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.

3. Application to Data on Desired Number of Children

Data from the 2013 Nigeria Demographic and Health Survey (NDHS) were analysed. Individual data were available for 10900 women aged 15-49 on their ideal family size. Respondents who did not have any living children were asked "If you could choose exactly the number of children to have in your lifetime, how many would that be?" This category of respondent is the focus of this article. The survey was designed to provide this information at state and zone levels, for both urban and rural areas. The hierarchical structure of the dataset as used in this study is therefore described as follows

Individual level: The woman is considered the lowest level and the unit of analysis in this study.

State level: Each woman belongs to one of the 37 distinct states.

Zonal level: Each state comes from one of the 6 geopolitical zones of the country.

Information collected from woman *i* from state *j* and zone *k* is given as y_{ijk}, (i=1,…, 10900), (*j*=1,…, 37), (*k*=1,…,6) is the desired number of children a woman desire to have in her lifetime.

3.1 Description of variables

The response variable (number of desired children in a lifetime) is a continuous variable measured at the individual level (lowest level) while the independent variables include; location of residence, highest education attained**,** religion, current age and the wealth index of the individual women. The multilevel model for the three levels is then written as

$$(Desired\ Number\ of\ children)_{ijk} = \gamma_{000} + \beta_1 Islam + \beta_2 No\ education + \beta_3 primary + \beta_4 Higher + \beta_5 Rural + \beta_6 Poor + \beta_7 Middle + \beta_8 Current\ age + v_{0k(zone)} + u_{0jk(state)} + e_{ijk(individual)}$$

The aim is to assess the extent to which the observed factors at various levels (individual, state and zone) affect the desired number of children where $e_{ijk(individual)}$ is nested within $u_{ojk(state)}$ which is further nested within $v_{0k(zone)}$. The variances σ^2_e , σ^2_{uo} and σ^2_{vo} represent the variances of random effects due to individual, state and zone respectively. The higher the value of σ^2_e the greater the degree of differences in the individual women desired number of children. Also, the higher the values of σ^2_{uo} and σ^2_{vo}, the greater the degrees of differences induced by state and zone clustering respectively and the higher the degree of similarity of the desired number of children by the women within the same state and zone respectively.

Since there is possibility of having some of the individual level (level 1) predictors to vary across either state or region, a post hoc test was carried out to know whether the variables vary and to know whether the Random coefficient model is better than the random intercept model using the Akaike Information Criterion (AIC) of Akaike (1974), given as $AIC = -2\log L + 2p$, where logL is the log likelihood and p is the number of parameters in the model. A model with lower AIC is preferred.

4. Results

Table 1 presents the intercept only model to determine the average number of children desired in a lifetime by women who have yet to have children. The average is approximately 5 children per woman which represents the grand mean. The intercept only model is given as;

$$(Desired\ \text{number of children})_{ijk} = \gamma_{000} = 5.187,$$

Table 1: Intercept only model.

Number of Children	Coeff.	Std. Error	P>|z|	95% Conf. Int	
Grand Mean	5.187	0.235	<0.001	5.141	5.233

Table 2 presents the Null Model with Random intercept without any variable at the three levels to calculate the intra-class correlation coefficient (that is, percentage variation in Number of desired children) between the states and geo-political zones. The method illustrated by Siddiqui *et al*. (1996) is employed to calculate the ICC; the intra-class correlations at the state and region level is given as

$$\rho_{state} = \frac{\sigma_{v0}{}^{2} + \sigma_{u0}{}^{2}}{\sigma_{v0}{}^{2} + \sigma_{u0}{}^{2} + \sigma_{e}{}^{2}} \qquad (9)$$

and

$$\rho_{zone} = \frac{\sigma_{v0}{}^{2}}{\sigma_{v0}{}^{2} + \sigma_{u0}{}^{2} + \sigma_{e}{}^{2}}, \qquad (10)$$

where $\sigma_{v0}{}^{2}$ is the variance of each zone form the grand mean, $\sigma_{u0}{}^{2}$ is the variance of each state mean within its zone and $\sigma_{e}{}^{2}$ is the variance form its state mean.

Table 2: The Null Model with Random intercept.

Random-effects Parameter	**Estimate**	**Std. Error**	**95% Confidence Interval**	
Zone: Standard deviation	1.103	0.333	0.610	1.994
State: Standard deviation	0.582	0.777	0.448	0.756
Individual: Standard deviation	2.150	0.146	2.121	2.178

Using (11) and (12), the intra-class correlations at state and zone levels are estimated as 0.2519 and 0.1971 indicating that 25 per cent of the variation in desired number of children comes from differences among the states and 19 per cent from differences among the zones.
To know whether the Random coefficient model is better than the random intercept model, the result showing the values of AIC of the two models are presented in Table 3.

Table 3: The values of AIC for random coefficient model and random intercept model.

Models	**AIC**
Random Coefficient Model	47017.824
Random Intercept Model	47061.166

As observed from Table 3, random coefficient model has the lower AIC value. A further discussion of effects of the observed factors on desired number of children is based on random coefficient Model.
Table 4 present the varying intercept model with the individual level predictors. As observed, a year increase in the current age of the women increases the desired number of children by 0.012. Women with no formal education are 0.851 times likely to desire more children compare to women who have secondary education while women with primary education are 0.416 times likely to desire more children but women with higher education desire for more

children reduced by 0.375 times compare to women with secondary education. Muslim are 0.918 times more likely to desire more children than Christian women.

Table 4: Varying Intercept model with individual level predictors

Factors	Desired Number of Children	Coefficient	Std. Error	P>|Z|	95% Confidence Interval	
Current Age	Current age	0.0118	0.0038	0.002*	0.0044	0.0192
Highest Level of Education Attained	Secondary (ref)					
	No Education	0.851	0.084	<0.001*	0.687	1.015
	Primary	0.416	0.072	<0.001*	0.275	0.557
	Higher	-0.375	0.6478	<0.001*	-0.502	-0.248
Religion	Christian (ref)					
	Islam	0.916	0.067	<0.001*	0.784	1.047
Location	Urban (ref)					
	Rural	0.087	0.053	0.098	-0.016	0.190
Wealth Index	Rich (ref)					
	Poor	0.410	0.072	<0.001*	0.294	0.519
	Middle	0.406	0.058	0.002*	0.004	0.019
	Constant	4.348	0.277	<0.001	3.806	4.890

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

Random-effects Parameter	Estimate	Std. Error	95% Confidence Interval	
Zone: Standard deviation	0.620	0.194	0.336	1.144
State: Standard deviation	0.416	0.058	0.317	0.547
Individual: Standard deviation	2.084	0.014	2.056	2.111

From the distribution of wealth, women categorized as poor and in the middle class are 0.410 and 0.406 more likely to desire more children compare to women that are in the rich category. Of all the factors considered, the place of residence of the women does not affect the desire for children.

5. Discussion and conclusion

The study was carried out on the desired number of children among women who are yet have any children using dataset from 2013 Nigeria Demographic and Health Survey (NDHS). For the study, a three-level multilevel regression model which account for hierarchical structure of the data was used. Results of Akaike Information Criterion (AIC) indicated that random intercept model with individual level predictor was a better in predicting the desired number of children. Results of finding showed location of residence does affect the desired number of children in Nigeria. Women with no education and primary education has higher desire to have more children while women with higher education, desire less number of children. The desire to have more children is prevalent among Muslim women than Christian women. There is high level of desire to have more by women whose wealth index is categorised as either poor or middle class.

We observed that, the desire to have more children is high by women with lower status in terms of education and wealth index. This is a pointer to the area in which the government needs to intensify effort. Education of the girl-child must be encouraged to enlighten them on the need to reduce the number of children given birth to. By doing these, there will be improvement in the socio-economic situation of the country. To this effect, the inclusion of the hierarchy of the data enriched the analyses carried out in this study rather than using the common regression analysis.

References

Akaike, H. (1974): A new look at the statistical model identification. *IEEE Transactions on Automatic Control*. **19,** 716-723.

Brown, L. R. (2011): *The new geopolitics of food Foreign Policy*. Retrieved from http://foreignpolicy.com/2011/04/25/the-new-geopolitics-of-foo

De Leeuw, J. and Kreft, I. G. G (1986): Random Coefficient Models for Multilevel Analysis. *Journal of Education Statistics*. **11**, 57-85.

Federal Ministry of Health (1998): *National Policy on population for Development, Unity, Progress and Self-reliance.* Federal Ministry of Health, Lagos, Nigeria.

Federal Republic of Nigeria (2004): Background Paper on *Employment and Poverty Alleviation in Nigeria* presented to the African Union Extraordinary Summit on Employment and Poverty Alleviation in Africa, Ouagadougou, 8th-9th September, 2004, Abuja, Federal Government Press.

Gill, J. (2003): *Hierarchical linear models.* In Kimberly Kempf-Leonard (Ed.), Encyclopedia of social measurement. Academic Press, New York.

Heather, W., Andrea, F., Mackay, J. C. and Rocchi, M. (2012): An Introduction to hierarchical linear modeling. *Tutorials in Quantitative Methods for*

Psychology. **8**(1), 52-69.

Hofmann, D. A. (1997): An overview of the logic and rationale of hierarchical linear models. *Journal of Management*. **23**, 723-744.

Hox, J. (2010): *Multilevel Analysis Techniques and Applications.* Second Edition. Published by Lawrence Erlbaum Associates Inc., Mahwah, New Jersey.

Lee R. and Miller T. (1990): Population policy and externalities to childbearing. *The Annals of the American Academy of Political and Social Sciences*. **510** (1), 17-32.

Longford, N. T. (1987): A fast scoring Algorithm for Maximum Likelihood Estimation in Unbalanced Mixed Models with Nested Random Effects. *Biometrika*. **78**, 817-827.

Longford, N. T. (1993): *Random Coefficient Models*. Oxford University Press, Oxford, UK.

National Population Commission (2004): *National policy on population for sustainable development. Abuja, Nigeria*: National Population Commission.

National population Commission (2000): *Nigeria Demographic Health survey 1999.* National population Commission and ORC/Macro, Calverton, Maryland.

National Population Commission (NPC) and ICF Macro. (2014): Nigeria Demographic and Health Survey 2013. National Population Commission and ICF Macro, Abuja, Nigeria.

Mbiti, J. S. (1970): *Concepts of God in Africa.* Praeger, New York.

McClelland, G. (1983): *Family-size desires and measures of demand*. Academic Press, New York.

Olusanya, P. O. (1971): Status differential in fertility attitudes of married women in two communities in western Nigeria. *Economic development and cultural change*. **19**, 641-651.

Thomson, E. (2001): Family Size and Preferences. *International Encyclopedia of the Social and Behavioral Sciences*. **2004**, 5347-5350.

Randenbush, S. W. and Bryk, A. S. (1986): A Hierarchical Model for Studying School Effects. *Sociology of Education*. **59**, 1-17

Randenbush, S. W. and Bryk, A. S. (2002): *Hierarchical Linear Models: Applications and Data Analysis Methods*. Second edition. Sage Publications, Newbury Park, CA.

Siddiqui, O., Hedeker, D., Flay, B. R. and Hu, F. B. (1996): Intraclass Correlation Estimates in a School based Smoking Prevention study: Outcome and Mediating Variables by Sex and Ethnicity. *American Journal of Epidemiology*. **144** (4), 425–433.