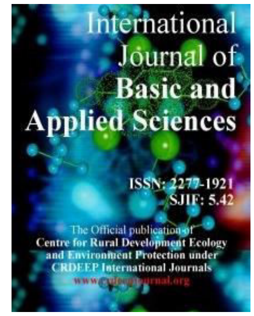
Vol. 9. No. 1. 2019

Contents available at:

www.crdeepjournal.org

International Journal of Basic and Applied Sciences (ISSN: 2277-1921)



Full Length Research Paper

Bayesian Multilevel Model on determinants of off-Farm Income Diversification in Assosa Woreda, Benishangul Gumuz, Ethiopia

Eshetu Mebrate[1]; O.Chandrasekhara Reddy[2]; Endale Alemayehu[2] and Mekdes Derbie[3]

[1.] *Department of Statistics, College of Natural Sciences, Assosa University, Ethiopia.*
[2.] *Department of Statistics, College of Science, EIT, Eritrea.*
[3.] *Department of Statistics, College of Natural Sciences, Ambo University, Ethiopia.*
[4.] *Department of Information Science, College of Computing and Informatics.*

ARTICLE INFORMATION

Corresponding Author:
O.Chandrasekhara Reddy

Article history:
Received: 10-12-2019
Accepted: 18-12-2019
Revised: 25-12-2019
Published: 30-12-2019

Key words:
Off-farming, Multilevel, Bayesian

ABSTRACT

Introduction: Agriculture is the foundation of Ethiopia's economy, which contributes 45% GDP, more than 80% of employment opportunities and over 90% of the foreign exchange earnings of the country, but unable to cover the food requirement of the country due to shocks and trends. The non-farm enterprise sector plays a vital role in enhancing the wellbeing of rural households as it provide them with income diversification opportunities. Objectives: The main objective of this research is to apply Bayesian multilevel model in identifying determinants of off-farm income diversification from the non-farm sector in Assosa Woreda, Benishangul Gumuz region. Methods: Step-by-step modeling has been used to analyze the data. Multilevel approach at both kebele and household level were considered. In Bayesian paradigm, the empty model, intercept model and intercept and slop model were employed. The convergence has been addressed using trace plot and autocorrelation. Results: The study indicated that Bayesian intercept model was found to be the best model. With this model, revealed that sex, age, family size, education, Level of current income, access to credit, food security, proximity to town and scarcity of food and income are found to be significant, indicating strong effects on off-farm income engagement. The results indicated that the variability in the off-farming in the household level was greater than that of kebele level. Conclusion: The intercept model is selected to be the best model among the under used model. In considering the multilevel cluster the variability of off-farming in household level is more diversified than that of kebele level. The convergence assumption has been tested by using trace plot and autocorrelation method.

Introduction

Agriculture is the foundation of Ethiopia's economy, which contributes 45% GDP, more than 80% of employment opportunities and over 90% of the foreign exchange earnings of the country, but unable to cover the food requirement of the country due to shocks and trends. Shocks are external events, such as erratic rain fall, flood, pests, disease, and market price fluctuation that adversely affect people's livelihood activities and trends are biophysical resources, demographic variables, and technical changes in production practices, economic condition, and educational status of households in a given area over time [1].

In recent years, the contribution of rural non-farm activities to household income diversification in developing world in general and sub-Saharan Africa in particular is increasing and contributes 30 to 45 percent of their income [2]. The non-farm enterprise sector plays a vital role in enhancing the wellbeing of rural households as it provide them with income diversification opportunities that helps in slowing down rural-urban migration, reducing poverty, and improving food security status [3]. [4]Argues that non-farm diversification is often a strategy that farm households use to moderate seasonal income variability and minimize the inherent risks associated with agriculture as a result of hostile agro-ecological factors.

Benishangul Gumuz Regional State (BGRS) is one of the poorest and most food insecure regions in the country. This is particularly, due to marginalization and isolation from development processes and initiatives [5]. In the region, agriculture accounts for about 93.2% of the people's livelihood, but its reward is poor due to labor-intensive rudimentary farming tools (shifting hoe cultivation practiced by indigenous people), high prevalence of crop pests, human disease and weeds (especially termite, malaria and striga respectively), erratic nature of rain fall, forest fire, poor infrastructure setting and extension services [6]. Despite increasing evidences on the potential contribution of non-farm income to economic wellbeing of rural households in the region, factors influencing their decision to diversify the activities are considerably overlooked. Therefore this study will investigate the determinants of non-farm income diversification in Assosa woreda; Benishangul Gumuz Regional State.

Benishangul Gumuz Region is remained one of the least developed regions in the country Ethiopia and majority of its people in the areas face the challenges of critical food shortages and less income diversification. There are many Factors contributing to this problem. All these interwoven activities and situations deteriorating various potentials and opportunities expressed above and poverty is still widespread and continuously increasing in the region. Therefore, this research will attempt to identify sample household's major non-farm diversified income generating activities and systematically examine the determining factors for its income diversification which was not specifically touched by other researchers in this research area.

To the extent of researcher knowledge, most studies upon off-farm income diversification have been done routinely based on the frequentist approach. Incorporating Bayesian approach for estimation can give accurate result compared to frequentist approach [7]. Therefore, the major gap of this study is applying Bayesian approach to identify predictors of off-farm income diversification and its associated factors and to point out the cross-Kebele variation taking households' Kebele as random effects.

Research Questions

1. Are there any association between off-farm income diversification and its covariate?
2. What are the factors that determine households' off--farm income diversification?
3. Does off-farm income diversification vary across households' Kebele?

The main objective of this research is to apply Bayesian multilevel model in identifying determinants of off-farm income diversification from the non- farm sector in Assosa Woreda, Benishangul Gumuz region.

The specific objectives of this research are to:

1. To investigate the association between off-farm income diversification and its covariate.
2. To examine factors determining households' off-farm income diversification.
3. To identify within and between households' Kebele variations.

Significance of the Study

Having clear picture and information on the livelihood strategies, food security status and their determinants in the study area can provide the basis for a detailed analysis on livelihood and food security in the country. A better understanding about non-farm income diversification behavior particularly in non agricultural sector will help in the design of policies that alleviates poverty, reduce vulnerability, and improve household well being. So, study will also provides direction for further research, extension, and development schemes that would benefit the farming population. Furthermore, the result may identify areas of intervention to alleviate poverty in general, food security in particular.

Materials and Methods

Description of the Study Area

Benishangul-Gumuz Regional State (BGRS): is one of the nine regional states established in 1995 by the new constitution of Ethiopia that created a federal system of governance. It is also recognized as one of the four 'emerging' or less developed regions in Ethiopia. Previously the southern part of BGRS belonged to Wollega while the area above the Abay River was under Gojjam province. The study woreda (Assosa): is one of the 21 woredas found in eastern parts of the Region, bordered by Kurmuk and Homesha in the north, Menge in the northeast, Oda Buldigilu in the east, Bambasi in the southeast, Mao-Komo special woreda in the south and by Sudan in the west. The 2007 national census reported that the total population of Assosa woreda is 87,366 of whom 23.25 percent was urban dwellers.

Sampling Procedures and Sample Size

A multi-stage sampling procedure will be employed in order to select sample households. In the first stage, out of the eight woredas (Assosa, Bambasi, Homosha, Kurmuk, Menge, Odabildagul and Sherkole) in Assosa Zone, Assosa woreda were the target of this research since its high population, accounts both settlers and indigenous people and residential areas of the researcher. In Assosa woreda there are 73 rural kebelles among which 35 of them are indigenous. And hence due to the existence of indigenous and settlers differing in culture and livelihood systems in the woreda, cluster sampling were used to identify woredas, where settler and indigenous people live.

In the second stage systematic random sampling were used to select sample population from each cluster. Systematic random sampling is the selection of every K^{th} element from a sampling frame, where K, is the sampling interval and K = total woreda in each cluster divided by sample size or K = N/n.

So, for, Indigenous population, K = 35/8 = 4.3 (four samples will be selected systematically as shown above). These are: Afaism, AtsetseAdirinunu, DabusAtimbaro and Rubalageda. In the case of Settlers, K = 38/8 = 4.3 (5 sample kebelles (Amba_1, Amba_10, Komshega_27, Menge_37 and Silga_24 will be selected).

In the third stage the desired number of households those who respond the interview questionnaires and Focus Group Discussions/FGD/ will be selected using the formula of sample size determination provided by YamaneTaro"s (1967) a simplified formula for sample sizes (n=N/1+N (e2),

Where;

N is the total population
n is sample size
e is the error margin.
thus,

$n=87366/(1+87366*(0.05)^2)$
$n=87366/219.415$
$n= 398$

Study Variables

Dependent variable:

The dependent variable considered for this study isoff-farm income diversification which a dichotomous random variable. Therefore, for this study, households who engaged in off-farm income were coded as 1 whereas households who were not engaged in off-farm income were coded as 0.

Explanatory variables

Explanatory variables considered in the study were selected based on some previous studies and those that are expected to be factors/determinants of off-farm income diversification. As suggested in the literature review, several variables that are associated with off-farm income diversification were considered as predictor variables. Therefore, those variables that are reviewed in the literature are listed below.

- Sex of house hold
- Age of house hold
- Education level of house hold
- Family size of house hold
- Marital Status
- Level of current income
- Food security of house hold
- Proximity to the nearest city
- Access to credit
- Feeding frequency of house hold above 15 years
- Feeding frequency of child
- Scarcity of food and income

Methods of data Analysis

There are different statistical analyses that researcher has used for data analysis. Under this study researcher used Bayesian multilevel logistic regression.

Bayesian Multilevel Analysis of Random Coefficient Model

Since the logistic regression mode can be changed to linear using the logit link function, similarly in the multilevel analogue, random coefficient logistic regression is based on linear models for the logit link function that include random effects for the groups or other higher level units. Consider explanatory variables which are potential explanations for the observed outcomes. We can denote these variables by $X_1, X_2, \dots, X_k$. The values of X_h(h=1,2,3,...,k) are can also be assigned in the usual way by X_{hij}, since some or all of these variables could be level one(households) variables, the success probability is not necessarily the same for all individuals in a given group(region). Therefore, the success probability depends on the individual as well as the group, and is denoted byπ_{ij}. Now consider a model with group specific regression of logit of the success probability $logit\ (\pi_{ij})$, on single level one explanatory variable X.

$$logit(\pi_{ij})=\left[\frac{\pi ij}{1-\pi ij}\right] = \beta_O + U_{Oj} + \sum_{h=1}^{K}\beta_h X_{hij} + \sum_{h=1}^{K} U_{hj} X_{hij} \ldots \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [11]$$

The expression $\sum_{h=1}^{K} U_{hj}X_{hij}$ can be considered as a random interaction between group and the explanatory variables. This model implies that the groups are characterized by two random effects: their intercepts and their slopes. It assumes that, for different groups the pairs of random effects (U_{Oj},U_{hj} h= 1,2,...,k, j=1,2,…11) are independent and identically distributed. The random intercept variance, $var(U_{Oj}) = \delta_O^2$,the random slope variance , $var(U_{1j})$=δ_1^2 and the covariance between the random effects, $cov(U_{Oj}, U_{1j})$ =δ_{01} are called variance components [8].

Likelihood Function

The likelihood function used in Bayesian approach is equivalent to that of the classical inference. The joint distribution of n independent Bernoulli trials is the product of each Bernoulli Densities, where the sum of independent and identically distributed Bernoulli trials has a Binomial distribution. Specifically, let $Y_{1j}, Y_{2j}, \ldots Y_{ij}$be independent Bernoulli trials with success probabilities$\pi_{1j}, \pi_{2j}, \dots \pi_{ij}$that is $Y_{ij} = 1$ (women contraceptive use) with probability π_{ij} and$Y_{ij} = 0$(not practicing off-farm income) with failure probability$1 - \pi_{ij}$, for i= 1,2,…,n and j=1,2,…11.Since, the trials are independent, the joint distribution of $Y_{1j}, Y_{2j}, \ldots Y_{ij}$is the product of n Bernoulli probabilities. The probability of success in logistic regression varies from one subject to another, depending on their covariates. Thus, the likelihood function is illustrated below as:

$$L(\pi_{ij}/\ yij) = \prod_{ij}(\pi ij)^{yij}\,(1 - \pi ij)^{1-yij} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots[12]$$

Where, π_{ij}represents the probability of the event for subject ij who has covariate vectorX_{ij}, $y_{ij} = 1$ indicates the presence (off-farm income) and Y_{ij}=0 the absence (not engaged off-farm income) of the event for the given subject. The probability of success in logistic regression can be defined as:

$$\pi_{ij}=\frac{exp\left(\beta_O + U_{Oj} + \sum_{h=1}^{K}\beta_h X_{hij} + \sum_{h=1}^{K} U_{hj} X_{hij}\right)}{1+exp\left(\beta_O + U_{Oj} + \sum_{h=1}^{K}\beta_h X_{hij} + \sum_{h=1}^{K} U_{hj} X_{hij}\right)}$$

Prior Distribution for random coefficient model

The prior distribution for the parameters for $\beta_0\ \beta_1 \ldots\ldots \beta_k$ and Ω_u has been denoted as follow:

$$f(\beta j \mid data) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left\{ \frac{-1}{2}\left(\frac{\beta_j - \mu_j}{\sigma_j}\right)^2 \right\}$$

$P(\Omega_u)$ $\propto$inverse − wishart(S_u^{-1}, v)denotes the inverse Wishart distribution with scale matrix $S_u$and degrees of freedom v. The parameter $\Omega_u$is the variance covariance matrices. Equivalently, information about a variance-covariance matrix is represented by means of a Wishart (S_u ,v) distribution placed on the precision matrix Ω_u^{-1}[9].

$$P(\Omega_u^{-1}) \propto wishart(S_u, v)$$

The Wishart distribution is the multivariate extension of the gamma distribution, although most statisticians use the Wishart distribution in the special case of integer degrees of freedom, in which case it simplifies to a multivariate generalization of the χ^2distribution.

Posterior Distribution for random coefficient model

.The posterior distribution is obtained as the product of the prior distribution of the parameters and the likelihood function. Therefore, using the above prior and likelihood function the full conditional posterior distribution for the parameters $\beta_0, \beta_1, \beta_2, \ldots \ldots \beta_K$ is given by:

$$f(\beta j \mid data) = \prod_{i=1}^{n}\left[\left(\frac{e^{\beta_0+\beta_1 X_{i1}+\beta_2 X_{i2}+\ldots+\beta_p X_{ip}}}{1+e^{\beta_0+\beta_1 X_{i1}+\beta_2 X_{i2}+\ldots+\beta_p X_{ip}}}\right)^{y_i}\left(1-\frac{e^{\beta_0+\beta_1 X_{i1}+\beta_2 X_{i2}+\ldots+\beta_p X_{ip}}}{1+e^{\beta_0+\beta_1 X_{i1}+\beta_2 X_{i2}+\ldots+\beta_p X_{ip}}}\right)^{1-y_i}\right] -$$

$$\times \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma_j^2}}\exp\left\{\frac{-1}{2}\left(\frac{\beta_j-\mu_j}{\sigma_j}\right)^2\right\}$$

[4]

Where h=1, 2,…K

And the full conditional distribution of the variance-covariance parameter Ω_u has been given as:

$P(\Omega_u / \beta_h U_{oj}, \quad Y_{ij})) \quad \propto \quad P(Y_{ij}/\beta_h, \Omega_u U_{oj},) P(U_{oj}/\Omega_u) P(\Omega_u)$……………………………[14]

Result and discussion

Descriptive statistics

This study was carried out to identify off-farm income diversification in Assosa woreda through analyzing the demographic and economic factors which were considered in similar studies conducted previously. In this study both descriptive and inferential analyses have been investigated for the purpose of identifying factor off-farm income diversification and food security nexus. Accordingly, the study used 398 households and the results are presented in two main parts. The first part of the result is the frequency distributions of all independent variables with their respective categories. The second part of the result is the chi-square test analysis (cross tabulation), with which the association between each explanatory variables and dependent variable (off-farm income). Consequently, the result obtained from descriptive analysis has shown number of households that are engaged in off-farm income. Therefore, the result indicated that out of 398 households considered in the analysis, 177(44.5%) households are engaged in off-farm income generating.

Test of heterogeneity proportions of off-farm income diversification

The two-level structure is used with the kebele as the second-level unit and the households' off-farm income as level one unit. This is based on the idea that there may be differences in households'off-farm income diversification between kebele that are not captured by the explanatory variables and hence may be regarded as unexplained variability within the set of all kebele [8]. Before attempting to multilevel analysis, one has to test the heterogeneity of households' off-farm income diversification among kebele of assosa woreda from which essential clues would be obtained for incorporating the random effects. Therefore, the Pearson chi-square for the proportion of households' off-farm income diversification across the kebele has been investigated in the table below. Consequently, as it can be observed in the Table 1, the Pearson Chi-square ($\chi^2 cal) = 53.672$ which is greater than15.507at 8 degree of freedom with P-value =7.99e-09 which is less than 0.05 level of significance, implying strong evidence of heterogeneity for off-farm income across kebele.

Test of Association

As mentioned earlier in methodology part in order to determine the association between off-farm income engagement and individual explanatory variables. Consequently, the result obtained in the Table 2 below clearly indicated that most of explanatory variables such as sex, age, education, marital status, Level of current income, feeding frequency households, Food security, Access to credit, Family size and Proximity to the town have significant association with at 5% level of significant. However, Feeding frequency of child, scarcity of food and income were statistically found to be non-significant at 5% level of significant.

Table .1 Chi-Square Tests of Heterogeneity of off-farm income between kebele.

Chi-square test Statistics	Value	D.fP-value
Pearson Chi-square	53.6728	7.99e-09
N of valid cases	398	

Table .2 Chi-square test of association between independent variable and dependent variable

Variable name	Degree freedom	Pearson's Chi-squared	p-value
sex	1	20.722	0.0132
age	2	33.894	0.0015
education	1	47.408	0.000035
Marital status	2	7.1999	0.02733
Family size	15	14.227	0.05084
Level of current income	3	22.006	0.044
Feed_frequency_child	2	4.6031	0.1001
Feed_frequency_household above 15 years	2	8.3292	0.01554
Access to credit	1	0.37797	0.0387
Food security	1	45.489	0.000026
Scarcity_food_and_income	1	0.67648,	0.4108
Proximity to the town	1	29.661	0.00172

Source: own computation.

Bayesian multilevel logistic regression model comparisons

Here the comparisons of Bayesian multilevel models such as multilevel empty model, random intercept model, and random coefficient model were conducted based on Deviance information criterion which is mostly used as model comparison in Bayesian analysis. Therefore, as it is shown in the Table 3 below the Bayesian random intercept model is appropriately fitting the off-farm income among kebeles of household data sets as compared to empty and random slope model.

Table 3. Bayesian multilevel model comparisons

Model comparison statistics	Empty model	Random intercept model	Random coefficient model
DIC	423.4853	241.0768	349.8605

Bayesian intercept model

The results of the Bayesian intercept model in table 4.3 revealed that sex, age, family size, education, Level of current income, access to credit, food security, proximity to town and scarcity of food and income are found to be significant, indicating strong effects on off-farm income engagement because their corresponding p-value is less 5% level significant(α). It is also observed that the odds of engaged off-farm income for young households are 27.32($OR = e^{3.3078}$) times more than old households engaged in off-farm income. The odds of engaged in off-farm income for female household is 0.93 ($OR = e^{-2.7271}$) less than male households. Similarly, the odds of engaged in off-farm income for literate households are 7.04 ($OR = e^{1.9511}$) times more than illiterate households. The odds of engaged in off-farm income for households whose level of current income high is 0.80($OR = e^{-1.5954}$)less than those whose level of current income is very high. And also the odds of engaged in off-farm income for households whose level of current income medium is 0.9 ($OR = e^{-2.2633}$) less than those those level of current income is very high. Furthermore, the odds of engaged in off-farm income for households whose level of current income low 0.9($OR = e^{-2.3078}$)less than those whose level of current income is very high. The odds of engaged in off-farm income for households is 0.0321 when average income of households increased by one unit.

The odds of engaged in off-farm income for households who can't get access to credit 0.79 ($OR = e^{-1.5421}$)less than households who can get access to credit. The odds of engaged in off-farm income for households who have food security are 3652.6($OR = e^{8.2032}$) times more than those who have no food security. The odds of engaged in off-farm income for households who live far from town is 0.69($OR = e^{-2.1398}$) less than households who live near to town. Moreover, the odds of engaged in off-farm income for households who have scarcity of food and income is 0.999($OR = e^{-7.0870}$)less than households who have no scarcity of food and income.

Table 4. Bayesian intercept model

Variable name		Fixed effect					
	Categories	$post.mean$	SD	Sd.error	2.5%	50%	97.5%
intercept	βo	0.2506	1.9446	0.0869	4.0022	-0.2333	3.4892
Sex	female	-2.7271	0.5403	0.0242	-3.7441	-2.7058	-1.599
	Ref(male)						---
Age	Young	3.3078	1.0167	0.0455	1.2676	3.2844	5.5523
	Adult	3.4413	1.0206	0.0456	1.5841	3.3871	5.6562
	Ref(old)						
Education	Literate	1.9511	0.5107	0.0228	0.9535	1.9372	2.9680
	Ref (illiterate)						
Marital status	Single	0.9122	1.0594	0.0474	-1.1965	0.8878	3.0758
	Divorced	0.6001	0.8973	0.0401	-1.0871	0.6095	2.2545
	Ref (married)						---
Family size	size	0.2418	0.0869	0.0039	0.4142	0.2403	0.0825
Level of income	High	-1.5954	0.7985	0.0357	-3.1919	-1.5369	-0.0361
	Medium	-2.2633	0.7016	0.0314	-3.7410	-2.2410	-0.9837
	Low	-2.3047	0.6963	0.0311	-3.7673	-2.2852	-1.1036
	Ref(very high)						
Feeding frequency HH	3 times	-0.2923	0.4909	0.0230	-1.2216	-0.2958	0.6957
	> 3 times	1.1860	2.3298	0.1042	-3.1044	1.0618	6.1970
	Ref(twice and less)						---
Feeding frequency of child	3 times	2.3725	1.4331	0.0641	-0.3597	2.2994	5.1914
	>3 times	2.0086	1.5373	0.0688	-0.9205	2.0142	4.8948
	Ref(twice and less)						
Food security	Yes	8.2032	1.3073	0.0585	6.2655	7.9159	11.4788
	Ref(No)						

Proximity to town	far	-2.1398	0.5094	0.0228	-3.0301	-2.1707	-0.9845
	Ref(near)						---
Scarcity food and income	No	-7.0870	1.0919	0.0488	-9.3797	-6.9628	-5.2836
	Ref(Yes)						
Access to credit	No	-1.5421	0.6317	0.0282	-2.7199	-1.5778	-0.3244
	Ref(Yes)						---
Random effect							
Kebele	var($\square^2_{\square}$)	2.289	1.981	0.08859	0.4957	1.124	7.503

Intra class correlation

Under this situation, researcher can usually interpret the variation of intercept $\square^2_{\square\square\square\square\square\square}$ between cluster(in this case Kebele) by considering the ICC which goes from 0 indicates perfect independence of residuals or the observations do not depend on cluster membership and 1 indicates perfect interdependence of residuals or the observations only vary between clusters[10].It is usually expressed as :-

$$\square\square\square = \frac{\square^2_{\square\square\square\square\square\square}}{\square^2_{\square\square\square\square\square\square} + \square^2_{\square}}$$

Where $\square^2_{\square\square\square\square\square\square}$ is the variance of the between cluster and $\square^2_{\square}$ the variance of the residual. But, in the context of logistic regression, there is no direct estimation or calculation of the residuals on the first level. Therefore, $\square^2_{\square}$ is the logistic distribution variance which always can be given the value $\frac{\square^2}{3}$ is 3.29. The intra kebele correlation coefficient for this study was estimated $\hat{\square} = \frac{2.289}{2.289+3.29}$=0.41.This indicated that about 41% of the total variability in off-farm income diversification due to the fact that differences across Kebele and the remaining unexplained 59% accounts the between households differences.

Checking Convergence

Checking the convergence of an MCMC algorithm would be a pre-condition issue for the exact estimation of the posterior distribution of interest. For this reason, both the length of the burn in period and the size of the MCMC output that was used for the posterior analysis could be specified by the user. The other most important problem for checking convergence is specification of the thinning interval, that is, the numbers of iterations researcher needs to discard until two successive observations become independent. Regarding this, Metropolis-Hasting algorithm methods was implemented for this study with 60000 iterations, 10000 burn-in terms discarded, and 50 thinning interval to make observations independent or low autocorrelation. Therefore, two different methods such as trace plot and density plot for monitoring convergence have been presented below.

Trace plots: This is the graph which would be plotted the number of iterations versus the generated values. In this graph convergence can be attained if all values are within a zone without strong periodicities (up and down periods). Therefore, the trace plots are all straight line which did not show up and down periods.

Density plot: the density plots are almost similar with normal plot. This is an indication that all posterior estimates were converged. The trace and density plots can be found in appendices (see appendix A).

Conclusion

This study was aimed to apply Bayesian multilevel model in identifying determinants of off-farm income diversification from the non- farm sector in Assosa Woreda, Benishangul Gumuz region.

In order to address the basic research questions, the researcher considered different statistical methods to select the appropriate sample size and corresponding model.

Three different models were considered and the Bayesian intercept model was found to be the best model. With this model, revealed that sex, age, family size, education, Level of current income, access to credit, food security, proximity to town and scarcity of food and income are found to be significant, indicating strong effects on off-farm income engagement. In identifying the within and between cluster correlation, the intra correlation methods was applied and indicted that the variability in the off-farming in the household level was greater than that of kebele level. The study also checked the convergence; especially the convergence in the random effects has been checked and fitted very well using trace plot and autocorrelation.

Competing interests

The authors declare that they have no competing interest

Funding

All the expenses for the accessing of the data and related were covered the expenses from the authors

Authors' contributions

This study was designed and compiled by EM as the principal investigator. The development of the basic research questions, identifying the problems and selecting appropriate statistical models have been done collaboratively by EM and EA. Edition of the overall progress of the work MD

Acknowledgements

We thank Benishangul Gumuz Agricultural office for supporting us all the necessary data and relevant materials.

References

1. Chadwick, D., Wei, J., Yan'an, T., Guanghui, Y., Qirong, S., & Qing, C. (2015). Improving manure nutrient management towards sustainable agricultural intensification in China. Agriculture, Ecosystems & Environment, 209, 34-46.

2. Haggblade, S., Hazell, P., & Reardon, T. (2005, June). The rural nonfarm economy: Pathway out of poverty or pathway in?. In International Food Policy Research Institute. The future of small farms. Proceedings of a research workshop, Wye, UK (pp. 151-178).

3. Haggblade, S., Hazell, P.B.R. and Reardon, T. 2007. *Transforming the Rural Non-farm Economy*, John Hopkins University Press, Baltimore, USA.

4. Jung, J. Y., Adelstein, B. D., & Ellis, S. R. (2000, July). Discriminability of prediction artifacts in a time-delayed virtual environment. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 44, No. 5, pp. 499-502). Sage CA: Los Angeles, CA: SAGE Publications.

5. Woldehanna, T. (2004, May). The Experiences of Measuring and Monitoring Poverty in Ethiopia. In Inaugural Meeting of the Poverty Analysis and Data Initiative (PADI). May (pp. 6-8).

6. Haddad, L., Alderman, H., Appleton, S., Song, L., &Yohannes, Y. (2003). Reducing child malnutrition: How far does income growth take us?. The World Bank Economic Review, 17(1), 107-131.

7. Congdon, P. (2007). Bayesian statistical modelling (Vol. 704). John Wiley & Sons.

8. Snijders, T. A., & Kenny, D. A. (1999). The social relations model for family data: A multilevel approach. Personal Relationships, 6(4), 471-486.

9. Gentleman, R., Hornik, K., & Parmigiani, G. (2009). Use R!.

10. Sommet, N., &Morselli, D. (2017). Keep calm and learn multilevel logistic modeling: A simplified three-step procedure using Stata, R, Mplus, and SPSS. International Review of Social Psychology, 30(1).

APPENDIX A

Fig A: Trace and density plot