International Journal of Applied Science and Mathematical Theory E- ISSN 2489-009X P-ISSN 2695-1908,
Vol 5. No. 3 2019www.iiardpub.org

# Modeling Road Traffic Crashes in Nigeria Using Negative Binomial and Generalized Poisson Regression Model (1960-2018)

**Enekelekele , Blessing and Etuk, Ette Harrison**
Department of Mathematics/Statistics
Faculty of Science,
Rivers State University,
NkpoluOrowokuoro, Rivers State,
Nigeria.

**ABSTRACT**
*Road traffic accidents is trending one of the challenges that have drawn the attention of researchers, statisticians, road users, families, business men, economy and the general public. Some of these challenges arising from Road traffic crashes include; death of loved one, fractures, loss means of livelihood and other various degrees of injury, therefore, it is against this background that this study was targeted at modeling Road traffic crashes in Nigeria using Poisson Regression model, while the specific objectives of the study include; assessing factors that contribute to high rate of road traffic crashes. Secondly, to fit an appropriate model to data on road traffic crashes in Nigeria and to examine the already estimated model used in determining and modeling road traffic crashes in Nigeria. The data for the study was sourced for and extracted from the Federal Road Safety online database from the year 1960-2018. The data extracted were used in the simulation of Poisson Regression models with the aid of statistical software called STATA version 14. The results obtained from the analysis revealed that the* Poisson regression could not capture over-dispersion. So, other forms of Poisson Regression models such as the Negative Binomial Regression and Generalized Negative Binomial Regression were also used in the estimation. However, the Negative Binomial Generalized Regression Model contains the least Akaike Information Criterion (AIC) based on the selections of the overall best-fitted model. *Hence, recommendations were made based on the results from the findings.*

**Key words** :*Modeling , Road , Traffic , Crashes*

## INTRODUCTION

### 1.1 Background to the Study

Every year more than a million people die as a result of road-related crashes worldwide, and some million sustained one injury to the other and according to World Health Organization (WHO, 2004), this might likely increase by 65 per cent in the next 20 years due to rapid increase in purchase of motor vehicle and usage in large developing countries.The emergence of globalization in the world today had led to a high demand in vehicular movement and other mobile devices. The movement of vehicle and other related traffic activities has become a key driver to economic growth; therefore, this call for the need for expansions of the transportation industry.

According to Atunbi (2009), Road traffic crashes are the leading causes of death among adolescent and younger people in their prime age. There has been a surge in the proportion, and an absolute number of traffic fatalities witnessed in several developing countries of the world while the industrial nations are experiencing decrease in downward trend in the occurrence of road traffic crashes by more than 20% (Emenike and Ogbole, 2008). They further opined that road traffic crashes situation in Nigeria has been alarming and particularly disturbing ever since the first auto crash was recorded. Sequel to the above disturbing facts, there is a need to develop a statistical approach in estimating road traffic crash with a view to establish safety technique to use in averting road traffic crashes in Nigeria. To develop a statistical approach in estimating road traffic crashes Cameron and Trivedi (1998) revealed that Poisson and Negative Binomial regression analysis are the best techniques to be used in modeling road traffic crashes data that appears in discreet or countable form. They further opined that Poisson and Negative Binomial regression are often applied in studying the occurrence of events that appears in countable form and also as a function of a set of predictor variable in an experimental and observational study in many disciplines including; Economics, Demography, Psychology, Biology and Medicine (Gardener *et al.*, 1995). Oppong and Assuah (2015), opined that Poisson and Negative Binomial regression models are often used as an alternative model to the Cox model for survival analysis when hazard rates are approximately constant during the observation period and when risks associated with an event under consideration is minimal (e.g., the incidence of road accidents) in such a case Poisson or Negative Binomial Regression model usually replaces with the Cox model. In such a case, there is the assumption that the Cox model cannot be quickly used in estimating aggregated data. This means that there are a lot of challenges associated with the use of the Cox model in modeling road traffic accident data. It is against this background that this study model road traffic crashes in Nigeria using Poisson, negative binomial and Generalized Negative Binomial regression models between 1960 to 2018.

## METHODOLOGY

This section was discussed under the following sub-headings; Source of data, model specification, estimation technique and procedures.

### 3.2 Source of Data

Data used in the study was sourced for and extracted from the Nigeria Federal Road safety On-line Statistical Database. The Variables comprises of annual data extracted from 1960 -2018, making it a total of 177 data points. The variables of interest for the study are classified into three: fatal, Minor and serious type of accident injury from road traffic crash casualties. Statistical package STATA version 14 was used in analyzing the data extracted for the study.

### 3.3 Model Specification

In line with the specific objective of this study, the models adopted for this study are Poisson regression, Generalized Poisson Regression, Negative Binomial Regression and Generalized Negative Binomial Regression. They are derived as shown below;

**Poisson Regression Model**

Poisson regression models are used basically when facing a problem whereby the outcome of the random process can take count values only, for example in a road traffic crash scenario it makes sense to assume that the number of road crashes that occur in daily, weekly, monthly or annually are considered as count data. According to Philip and Sebastian (2015), one of the distributions that satisfies such criterion comes from the family of exponential distribution which is the Poisson distribution.Let Y be a random variable (the rate at which the road traffic crashes occurs) and let $y_i,\ i=1,\ 2,\ 3$ be the outcomes of the road traffic crash as an event. The variable $Y$ is said to follow Poisson distribution with parameter $\mu > 0$ if the probability function is given by

$$P(Y=n)=\frac{e^{-u}\mu^{n}}{n!}$$

Where $n=1, 2, 3$ is the number of occurrences of an event and $\mu$ is defined as $\mu = E[Y]$. One of the useful properties of the Poisson distribution is that the variance depends on the mean and also the variance is equal to the mean. The generalized linear model can be stated as thus:

$$\eta_i = \beta_0 + \beta_1 X_{i1} + \ldots\ldots.. + \beta_k X_{ik} \qquad (3.1)$$

The two link functions are stated statistically as follows:

The first link function describes how the mean $E(Y_i) = \mu_i$ which depends on the linear predictor

$$\vartheta(\mu_i) = \eta_I \qquad (3.2)$$

The second link functions describes how the variance, $Var(Y_i)$ depends on the mean

$$Var(Y_i) = \phi var(\mu) \qquad (3.3)$$

whereas the dispersion parameter $\phi$ is a constant , supposing $Y_i$ is a Poisson distribution

Then; $Y_i \sim$ Poisson $(\lambda_i)$,

$$E(Y_i) = \lambda_i,\ var(Y_i) = \lambda_I \qquad (3.4)$$

Therefore, our variance function;

$$Var(\mu_i) = \mu \qquad (3.5)$$

and the link function must map from $(0,\infty)$. A natural log of the function given as

$$\ell(\mu_i) = \log_e(\mu_i) \qquad (3.6)$$

The generalized linear regression model (GLM), according to Nwankwo and Nwaigwe (2016) has to do with allowing the linear model to be related to the response variable through linked function. The link function here is the function that links between the linear model in a design matrix and the Poisson distribution function. Supposing a linear regression model given as thus:

$Y_i = \beta_i X_i \ x\varepsilon_i$ (3.7)
If XE IR$^n$, is a vector of independent variables

$$Y = X\beta x\varepsilon \quad (3.8)$$

Where X is an nx(k + 1) vector of independent variables of predictors, and a column of I's β is a (k + 1) by 1 vector of unknown parameters and ε is an n x 1 vector of random error terms with mean zero. Therefore,

$E(^Y/_X) = X\beta$ (3.9)

Recall that the Generalized linear models, where the link function and its transport Y as;

$G(y) = \log e\ (y)$ (3.10)

Therefore, this can be written in more concise form as;

$\text{Log}_e\ E(^Y/_X) = X\beta$ (3.11)

Thus, given a poisson regression model with parameter β and its input vector X, the predicted mean of the associated poisson distribution is given as;

$$E(^Y/_X) = e^{XB} \quad (3.12)$$

Suppose $Y_i$ are independent observation with a corresponding values $X_i$ as the predictor variable, then the parameter β can be estimated using the maximum likelihood method. According to Nwanko and Nwaigwe, (2016) the model expressed in equation (3.12) can be estimated by numerical methods and this is done using the logarithmic transformation of the conditional expectation of the dependent and independent variables. He further explained that the probability surface of the maximum – likelihood estimation of Poisson regression models are always convex form such that Newton-Raphson of the gradients –based methods are use as an appropriate estimation techniques.Therefore, suppose $Y_i$ is a random variable and it takes non-negative values such that $i$ = 1, 2 …… n, where n is the number of observations. Since $y_i$ follows a Poisson distribution, therefore the probability mass function (PMF) is as thus

$$P(Y = y_i) = \frac{\lambda_i^{y_i} e^{-\lambda}}{y_i!} \quad , \quad y_i = 0, 1, 2 \quad (3.13)$$

With mean and variance as

$$E(y_i) = Var(y_i) = \lambda \quad (3.14)$$

Where the conditional mean (predicted mean) of the Poisson distribution as given in equation (3.12) above specified as;

$$E(^y/_x) = e^{x\beta} = \lambda = E(y_i) \quad (3.15)$$

Where it is the value of the explanatory variable $\beta = (\beta_1, \beta_2, \ldots\ldots\ldots \beta_k)$ are unknown K – dimensional vector of regression parameters and the mean of the predicted Poisson distribution is given as $E(^Y/_X)$ and its corresponding variance of $Y_i$ as $var(^Y/_X)$.

**Generalized Poisson Regression Model**

One of the advantages of using the generalized Poisson regression model is that it can be fitted to over-dispersion model i.e. where $Var(y_i) > E(y_i)$ as well as under-dispersion, $Var(y_i) < E(y_i)$ Nwankwo and Nwaigwe, 2016). Famoye (1993) Wang and Famoye (1997) further suggested that when $y_i$ is a count response variable and its follows a Generalized Poisson distribution, the probability density function given that $y_i$ $i = 1, 2$ ….. n, then

$$f(y_i) = \rho(y_i = y_i) = \left[\frac{\mu_i}{1+\alpha_i\mu_i}\right]^{y_i} \frac{(1+\alpha y_i)^{y_{i-1}}}{y_i!} \exp\left[\frac{V_i(1+\alpha y_i)}{(1+\alpha\mu_i)}\right], \; y_i = 0, 1, \qquad (3.16)$$

Where, mean $E(y_i) = \mu_i$ and variances $var(y_i) = (\mu+\alpha,\mu_i)^2$ and $\mu$ is refers to as the dispersion parameter. Nwakwo and Nwaigwe,(2016) revealed that generalized Poisson distribution is a natural extension of the Poisson distribution. When $\alpha = 0$, the model in equation (3.21) reduces to the Poisson (as in equation 3.13), whereby $Var(y_i) = E(y_i)$. When $\alpha > 0$, it means the variance $Var(y_i)$ of the distribution represents count data with over-dispersion if $\alpha < 0$, it means the variance is less than the expectation ie $Var(y_i) < E(y_i)$ which simply means that the distribution represents count data with under-dispersion. Supposing it is assumed that the mean of the fitted value is multiplication i.e. $E(y_i/X) = \mu_i = e_i \exp(x_i\beta)$

Where $e_i$ denotes a measure of exposure, Similarly $x_i \alpha pX1$ vector of explanatory variables and $\beta\alpha pX1$ vector of regression parameters (Nwonkwo and Nwaigwe, 2016).

**Negative Binomial Regression (NBR)**

According to Shaw-Pin (1993), negative binomial distribution is used to deal with the problem of over-dispersion in count data. Over-dispersion occurs when there is the presence of statistical variability in a data set. A situation in which theoretical population mean of a model is approximately the same as the sample mean. It can be further explained that this occurs when the observed variance is higher than the variance of a theoretical model, then over-dispersion is said to occur. On the other hand, under-dispersion simply means that there was less variation in the data than predicted. Over-dispersion is a very common characteristic in applied data analysis because in practice, populations are frequently heterogeneous (non-uniform) opposed to the assumptions implicit within widely used simple parametric models. The Negative Binomial regression model used in this study was specify as thus:

$$P(Y_i = y_i) = \frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{\Gamma(y_i+1)\Gamma\left(\frac{1}{\alpha}\right)} \left(\frac{1}{1+\alpha\mu_i}\right)^{\frac{1}{\alpha}} \left(\frac{\alpha\mu_i}{1+\alpha\mu_i}\right)^{y_i}, \; y_i, i = 0,1,2,3........ \qquad 3.17$$

Where the mean is given by

$$\mu_i = E(y_i) = V_i[e^{\sum_i \alpha x_{ij}\beta_j}] \quad i = 1,2,3,.........n \qquad 3.18$$

The variance of $y_i$ is given as thus:

$$Var(y_i) = \mu + \alpha\mu_i^2 \qquad 3.19$$

Where $\alpha > 0$ the model would be referred to as dispersion parameter. From Eq. (3.25) one can see that this model allows the variance to exc*eed* the mean. Also, the Poisson regression model can be regarded a*s* a limiting model of the negative binomial regression model as $\alpha$ approaches 0.

**Generalized Negative Binomial Regression (NBR)**

The generalized Negative Binomial Regression (NBR) distribution given as thus:

For 1 > a > 0 and $|\alpha\beta| < 1$ we define the generalized negative binomial distribution by

$$b_\beta(x,n,\alpha) = \frac{n\Gamma[n+B_x]}{X!\Gamma[n+\beta x - x + 1]}\alpha^x(1-\alpha)^{n+\beta x - x}, n > 0; x = 0,1,2,......... \qquad 3.41$$

Such that $b_\beta(x,n,\alpha) = 0$ for $x \leq \underline{m}$ if $n + \beta m < 0$

Similarly, the estimation of Generalized Negative Binomial Regression is done using the Maximum Likelihood method of Parameters estimation

**3.4 Model Selection Test**

Model selection shall be done using two criteria which include: the Akaike information criteria (AIC) and Bayesian information criteria (BIC). It is defined as below in the two models:

AIC(n) = $\frac{-2}{n}[L-K]$ and AIC(1) = -2[L – K]

Where K is the number of predictors including the intercept, while AI(1) is usually an output in the statistical software applications. L is the maximized value of the likelihood function for the model. Similarly, Bayesian information criteria (BIC) according to Hube (2014) have three forms of mutations and they include as it is defined in (Schwart, 1978).

BIC(R) = D – (df) In (n)

B1C(L) = -2L + Kin(n)

BIC(Q) = $\frac{-2}{n}$(L – Kin (K))

Where;df is the residual degree of freedom, BIC (L) is given as SC in SAS and BIC in other software andL represents the log likelihood

**RESULTS**

This section focuses on the presentation of the results of the estimation of the model specified in chapter three fitting it to road traffic crash data extracted from Nigeria Federal Road Safety On-line Statistical Database.

**Table 4.1: Summary Descriptive Statistics of Accident Data Extracted from Federal Road Safety Online Statistical Data Based**

| Variable | Observations | Mean | Std-Dev | Min | Max |
|---|---|---|---|---|---|
| Fatal | 59 | 3284.6 | 1699.4 | 129 | 6986 |
| Serious | 59 | 8554.0 | 5180.6 | 1520 | 17352 |
| Minor | 59 | 7478.2 | 8201.7 | 841 | 19624 |
| No. of Incid | 59 | 21562.8 | 17.176 | 7771 | 41165 |

***Source: Researcher's Computation, 2018 using STATA Version 14.***

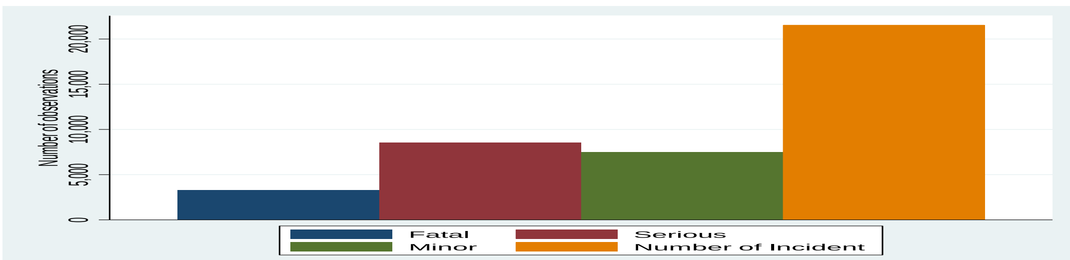


**Fig. 4.1: The Histogram Representation of Summary Descriptive Statistics of Accident Data.**

**Table 4.2: Multi-Collinearity Test Results**

| Variables | VIF | $\frac{1}{VIF}$ |
|---|---|---|
| Serious | 3.83 | 0.260943 |
| Minor | 2.050 | 0.262753 |
| Fatal | 1.02 | 0.977081 |
| Mean VIF | 2.89 | |

***Source: Researcher's Computation, 2019 using STATA Version 14.***

**Table 4.3: Estimation Results of Poisson Regression, Poisson Generalized Regression, and Negative Binomial Regression**

| Model | Indicator | Co-eff | Std Error | Z | P>/z/ |
|---|---|---|---|---|---|
| Poisson Regression | Constant | 9.209 | 0.004 | 2381.12 | 0.000 |
| | Fatal | 0.0001 | 5.20e-07 | 118.62 | 0.000 |
| | Serious | 0.0012 | 6.22e-07 | 186.12 | 0.000 |
| | Minor | -0.00006 | 3.50e-07 | -170.09 | 0.000 |
| | | | P-value | | |
| | Deviance | 127396.5 | 0.000 | | |
| | Chi-square | 130361.1 | 0.000 | | |
| | AIC | 128097.1 | | | |
| | BIC | 128105.4 | | | |
| Negative Binomial Regression | Constant | 9.178 | 0.1626 | 56.44 | 0.000 |
| | Fatal | 0.000872 | 0.00025 | 3.18 | 0.001 |
| | Serious | 0.000107 | 0.00003 | 4.21 | 0.000 |
| | Minor | -0.000057 | 0.00002 | -3.75 | 0.000 |
| | Model fitness | | | | |
| | AIC | 1214.342 | | | |
| | BIC | 1224.729 | | | |
| Generalized Negative Binomial Regression | Constant | 1.133261 | 0.1085 | 10.44 | 0.000 |
| | Fatal | 0.000144 | 0.0002 | 7.40 | 0.000 |
| | Serious | 0.000078 | 0.00002 | 3.99 | 0.000 |
| | Minor | -0.00001 | 0.00002 | -0.67 | 0.505 |
| | Model fitness | | | | |
| | AIC | 1205.164 | | | |
| | BIC | 1217.63 | | | |

***Source: Researcher's Computation, 2019 using STATA Version 14.***

**Table 4.4: Model Selection and -2log-Likelihood (-2ll).**

| Model(s) | AIC | BIC | -2ll | Overall Best fitted |
|---|---|---|---|---|
| Poisson | 128097.1 | 128097.1 | -91901.04 | |
| Negative Binomial Regression | 1214.342 | 1224.729 | -613.5994 | |
| Generalized Negative Binomial Regression | 1205.164 | 1217.63 | -611.2323 | 1205.164*** |

***Source: Researcher's computation, 2018 using STATA version 14.***

Overall Best fitted Model: Generalized Negative Binomial

## 5.1 Discussion of Results

The annual road traffic crash data collected was fitted with a family of count models such as: Poisson Regression, Negative Binomial Regression and Generalized Negative Binomial Regression).The data spanned from 1960 –2018.The summary descriptive statistic as shown in table 4.1 revealed that the total observations were 177 while the mean of those that have fatal accident was (3284.6) and its corresponding standard deviation was (1699.4). Similarly, those that have serious accident have the mean value (8554.0) with a standard deviation (5180.6) and minor accidence has a mean value of (7478.2) and its corresponding standard deviation was (8201.7). However, the total number of the incident recorded between the various degrees of injury has the mean value of (21562.8) with a standard deviation (17.176). From the results obtained so far, we can conclude that the high mean rated indicator of the various degree of injury related to the road traffic crashes that occur within the year under investigation was the total number of the incident recorded, followed by serious, fatal and minor injury respectively.

The result of the summary descriptive statistic was synonymous to Shaw-Pin (1993) findings. Shaw-Pin (1993), investigated the relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial re*g*ressions. The result obtained in Shaw-Pin (1993) investigation shows that truck related accident of high frequencies with various degree of serious injury across road sections were those that were victims. Although, there were a little variations in the two results and the variations in the two results could be attributed to methodology, geographical location and duration of the two studies. Figure 4.1 is the statistical representation of the various degree of injury sustained in an accident case which is shown on the histogram and the result indicates total number of incident represent the highest proportion which accounts for about 60.0% of the entire data used in the study followed by serious cases , fatal and minor respectively. In another development, table 4.2 shows the results obtained from the test for multi-collinearity between the variables (indicators) under consideration, this was done to verify whether there exist a perfect linear relationship between the predictors and the explanatory variable such that the estimates in the regression model can be uniquely computed.

From the results obtained the variance inflated factor(VIF) of all the variables have the same value (2.89) and according to the rule of thumb, a variable whose variance inflated factor(VIF) are greater than 10 may not merit further statistical investigation, however, the result is different here. Similarly, the tolerance value ($\frac{1}{VIF}$) value for serious injury (0.260943), Minor injury (0.262753), fatal injury (0.977081) and all the value of ($\frac{1}{VIF}$) obtained were lower than 0.1 comparatively to the VIF value 10. This simply means that all the variables can be included in linear combination with other independent variables and also in subsequent analyses, and modeling of Poisson regression as it was suggested in Oppong and Assuah (2015), examination in comparative assessment of Poisson and negative binomial regressions as best models for road count data and Also in Philip and Sebastian (2015) application of Poisson Regression on traffic Safety, the study presents a model that explains the traffic fatality by exploring the Poisson regression model using two types of explanatory variables – referred to as internal and external factors. The results obtained in Oppong and Assuah (2015), and Philip and Sebastian (2015) for the test for multi-collinearity seem to agree the present findings.

The results in table 4.3 shows that those who sustained fatal injuries have positive co-efficient (0.0001) at the 5% level of significance and this simply means that 1% increase in the cases of those who sustained fatal injuries in road traffic crashes in Nigeria during the period under investigation may lead to (0.01%) positive effect on the number of observed incidence of the cases of road traffic crashes in Nigeria. The z-statistics value (118.62) of the co-efficient of those that have fatal injuries is greater than 2 by the rule of the thumb, showing that the cases of victims of road traffic crashes having fatal injuries has a significant effect on number of observed incidence of the cases of road traffic crashes in Nigeria.

In another development, the result also revealed that those who sustained serious injuries has positive co-efficient (0.0012) at the 5% level of significance and this simply means that 1% increase in the cases of those who sustained serious injuries in road traffic crashes in Nigeria during the period under investigation may lead to (0.12%) positive effect on the number of observed incidence of the cases of road traffic crashes in Nigeria. The z-statistics value (186.12) of the co-efficient of those that have serious injuries is greater than 2 by the rule of the thumb, showing that the cases of victims of road traffic crashes having serious injuries has a significant effect on number of observed incidence of the cases of road traffic crashes in Nigeria.

Similarly, it was also confirmed that the estimate of those who sustained minor injuries has negative co-efficient (-0.00006) at the 5% level of significance which simply means that 1% increase in the cases of those who sustained minor injuries in road traffic crashes in Nigeria during the period under investigation may leads to (-0.006%) negative effect on the number of observed incidence of the cases of road traffic crashes in Nigeria. The z-statistics value (-170.09) of the co-efficient of those that have minor injuries is less than 2 by the rule of the thumb, showing that the cases of victims of road traffic crashes having minor injuries has a negative significant effect on number of observed incidence of the cases of road traffic crashes in Nigeria.

Furthermore, the value of deviance (127396.5) and Pearson chi-square (130361.1) greater than 1, so we can conclude that in estimating road traffic crashes using Poisson regression models the degree of injuries sustained from the various cases of road traffic crashes in Nigeria suffered from over-dispersion. According to Ayunanda, *et al* (2013), Over-dispersion usually leads models into producing biased parameter estimates. The consequences of over-dispersion is that it makes the value of the estimated standard error to be wrong as the mean value of the model is not to the variance as one of the precondition for Poisson regression models. Subsequently, may lead to errors in drawing inference about the parameters of the model.

To overcome these challenges of modeling counts data, we then resort to the use of negative binomial Poisson regression and the generalized negative binomial regression, since both models can accommodate and capture dispersion parameter in modeling counts data. However, the result in the estimation using Poisson regression model confirmed the assertion of Oluwaseyi and Gbadamosi (2017) in their investigation of road traffic Crashes in Nigeria: Causes and Consequences. In Oluwaseyi and Gbadamosi (2017) study, it was revealed that Motor vehicle crashes are the leading causes of death in adolescent and people in their prime age. It was also confirmed that there has been an expansion in the proportion and absolute number of traffic fatalities witnessed in a number of developing countries while the industrial nations are witnessing descending trend in the occurrence of accident by more than 20%. Also, this result corroborates Philip and Sebastian (2015) findings in the application of Poisson Regression on traffic Safety. However, Philip and Sebastian (2015) study presents a model that explains the traffic fatality by exploring the Poisson regression model using two types of explanatory variables – referred to as internal and external factors.

In Philip and Sebastian (2015) study, it was revealed that, the variables economic development, traffic exposure and demographic development significantly contribute to explain the long term cyclical trends, showing that traffic fatality is a complex multivariate system where no single variable can solely explain its dynamics. The external factor seasonal trend has the most impact of the examined external factors and explains the yearly cyclical pattern by itself. The model presented in this study shows high explanatory power and overall good fit to fatality data, making it a promising tool for statistical analysis of factors contributing to fatality.

In like manner, the estimation of the negative Binomial Regression shows that all parameters are significance level of 5% and it can be seen that the p-values of all parameters are smaller than 0.05. So, the negative binomial regression model revealed all the co-efficient of the various degrees of injuries to be positive except the case of minor injury. This confirmed that in estimating the co-efficient of negative binomial regression, 1% increase in fatal and serious cases of injuries may leads to 0.0872 % and 0.0107% respectively increase in the number of observed incidence of the cases of road traffic crashes in Nigeria. Also, their z-statistics values are (3.18 and 4.21) which are greater than 2 by the rule of the thumb, this confirming that fatal and serious cases of injuries have great significant effect on number of observed incidence of the cases of road traffic crashes in Nigeria. However, the case of those that had minor injury has negative co-efficient (-0.000057) at the 5% level of significance which simply means that 1% increase in the cases of those who sustained minor injuries in road traffic crashes in Nigeria during the period under investigation may lead to (-0.006%) negative effect on the number of observed incidence of the cases of road traffic crashes in Nigeria. The z-statistics value (-3.75) of the co-efficient of those that have minor injuries is less than 2 by the rule of the thumb, showing that the cases of victims of road traffic crashes having minor injuries has a negative significant effect on number of observed incidence of the cases of road traffic crashes in Nigeria.

The result obtained here using negative binomial regression in modeling the number of observed incidence of the cases of road traffic crashes in Nigeria corroborates Shaw-Pin (1993) study in investigating the relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. In Shaw-Pin's (1993) investigation, it was found that Poisson regression model should be used as an initial model for developing the relationship between all categories of accidence. Furthermore, it was also opined that if the over-dispersion of accident data is found to be moderate or high, both the negative binomial (NB) and Zero Inflated Poisson regression models could be explored. Also, the result obtained here further confirmed Shaw-Pin's (1993) assertion in investigation about the relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. Shaw-Pin (1993) asserted that under the maximum likelihood method, the estimated regression parameters from all the three models were quite consistent and no particular model outperforms the other two models in terms of the estimated relative frequencies of truck related accident involvements across road sections. In this present study, the result reveal the same assertion as the estimated regression parameters from all the three models were quite consistent with the same positive and negative sign for fatal, serious and minor injuries sustained in road traffic crashes in Nigeria within the period under investigation.
Also, the result obtained in this study was in line with Oppong and Assuah (2015) study. Oppong and Assuah (2015) examine comparative assessment of Poisson and negative binomial regressions as best models for road count data and found out that negative binomial regression model best fits road accidents' data significantly as compared with the poison regression model. Although, the little variation between the two results were due to geographical location and nature of the data counts, Oppong and Assuah (2015) study was carried out in Ghana using weekly road accident data whereas this present study uses annual data extracted from Nigeria federal road safety online statistical data base.

In another development , the estimation of the generalized negative Binomial Regression shows that all parameters are significance level of 5% and it can be seen that the p-values of all parameters are smaller than 0.05. So, the negative binomial regression model revealed all the co-efficient of the various degrees of injuries to be positive except the case of minor injury. This confirmed that in estimating the co-efficient of negative binomial regression, 1% increase in fatal and serious cases of injuries may leads to 0.0144% and 0.0078% respectively increase in the number of observed incidence of the cases of road traffic crashes in Nigeria. Also, their z-statistics values are (7.40and 3.99) which are greater than 2 by the rule of the thumb, this confirming that fatal and serious cases of injuries have great significant effect on number of observed incidence of the cases of road traffic crashes in Nigeria.

However, the case of those that had minor injury has negative co-efficient (-0.00001) at the 5% level of significance which simply means that 1% increase in the cases of those who sustained minor injuries in road traffic crashes in Nigeria during the period under investigation may leads to (-0.001%) negative effect on the number of observed incidence of the cases of road traffic crashes in Nigeria. The z-statistics value (-0.67) of the co-efficient of those that have minor injuries is less than 2 by the rule of the thumb, showing that the cases of victims of road traffic crashes having minor injuries has a negative significant effect on number of observed incidence of the cases of road traffic crashes in Nigeria.

### 6.2 Conclusion

The research was aimed at first, to examine the significance of the occurrence and incidence of road traffic crashes in Nigeria and secondly to assess the factors that may likely contribute to road traffic crashes in Nigeria.

The Poisson regression, Poisson Generalized Regression, Negative Binomial regression, and Generalized Negative Binomial Regression for the occurrence of road traffic crashes in Nigeria were considered.

### 6.3 Recommendations

Sequel to the results of the findings, the following recommendations were made;

1. Road traffic fatalities and injuries are, to a great extent, preventable, since the risk of incurring injury in an accident is largely predictable and there are many counter measures, proven to be effective. The most effective way to reduce fatalities and injuries would be through an integrated approach involving close collaboration of many sectors.
2. Progress is being made in many parts of the world where multi-Sectorial strategic plans are leading to reductions in the number of road accidental fatalities and injuries. Such strategies focus on four key factors that contribute to the risk of occurrence of a road accident, they are; exposure, behavioral factors, road environment, and vehicle factors. Perhaps the least used of all road safety intervention strategies are those that aim to reduce exposure to risk.
3. Risk in road traffic arises out of a need to travel – to have access to work or for education or leisure pursuits. Therefore, there is a need to promote not only regional economies in such a way that reduces the need for long-distance travel but also self sufficient compact townships which would reduce the need for short-distance travel within the cities.
4. The problem of road accidents in Nigeria also gets aggravated due to mixed nature of road traffic on its roads – with pedestrians, bicycles, mopeds, scooters, motorcycles, auto-rickshaws, taxis, vans, cars, trucks, and buses sharing the same road space.
5. In other words, the same road network is used by different categories of motorized and non-motorized vehicles, of varying width and speed. To reduce the exposure to risk, there is a need not only to segregate fast moving vehicles from slow moving vehicles or heavy vehicles from light vehicles but also enforce speed limit on fast moving vehicles. Road accidents and related injuries and fatalities are highly dependent on the speed of motor vehicles. Empirical

evidences suggest that an average increase in speed of 1 Km/h is associated with a 3% higher risk of a crash involving an injury.

## REFERENCES


Abdel-Aty, M.A. Radwan, A.E., (2000). Modeling traffic accident occurrence and involvement. *Accident Analysis and Prevention*, 32(5), 633–642

Afolabi, O. J &Kolawole, T. G(2017). Road Traffic Crashes In Nigeria: Causes And Consequences. *Causes and consequences. Transport & Logistics: the International Journal* 17(42), 2406-1069

Afolabi, J. O. (2009). “Topical Issues on Road Safety in Nigeria”. *International Journal* 17(42), 2406-1069

Aguero-Valverde, J., Jovanis, P.P., (2008). Analysis of road crash frequency with spatial models. Transportation Research Record: *Journal of the Transportation Research Board* 2061, 55-63.

Atunbi, A. O., (2009)“Urban Transportation,” An Appraisal of Features and Problems in Nigeria Society. *International Journal of Geography and Regional Planning*, 1(2),12-13

Barrett, S.D. (1991). *Transport Policy in Ireland in the 1990’s*. Ireland: Gill and MacMillan.

Bingham, C.R., Shope, J.T., Zhu, J., (2008). Substance-involved driving: predicting driving after using alcohol, marijuana, and other drugs. *Traffic Inj. Prev*. 9 (6), 515–526.

Black, J. (2002). *Oxford dictionary of economics (2nd edi.)*. New York: Oxford University Press Inc.

Brestow, N.E. (1990). Testing of hypotheses in over – dispersed Poisson regression and other quasi – likelihood models. *Journal of the American Statistical Association*, 85(1990), 565 – 571.

Brüde, U. (2005): What is happening to the number of fatalities in road accidents? A model for forecasts and continuous monitoring of development up to the year 2000. Accident Analysis and Prevention, 27, 405-410, 1995.

Cameron, A. C., Trivedi P. K. (1998). *Regression Analysis of Count Data*. U.K, Cambridge University Press, Cambridge.

Cameron,A.C& Trivedi,P.K(1998). *Regressionanalysisofcountdata*.Cambridge: Cambridge Universitypress.

Carson, J. & Mannering, F. (2001). The effect of ice warning signs on ice-accident frequencies and severities. *Accidents Analysis and Prevention*, 33, 99-109.

Chike, H. N & Godwin, I.(2016). Statistical model of road traffic crashes data in Anambra State, Nigeria: a Poisson regression approach. *European Journal of Statistics and Probability*, 4(1), 41-60.

Christian, P., Katz, J., Wu, L., Kimbrough-Pradhan, E., Khatry, S.K., LeClerq, S.C. (2008). Risk factors for pregnancy-related mortality: a prospective study in rural Nepal. *Public Health*, 122(2), 161-72.

Cook, J.D. (2009). Notes on the negative binomial distribution. (accessed on Feb. 15th, 2018). www.johndcook.com/negative_binomial.pdf.

Deebom Z.D and Essi I.D (2017). Modeling price Volatility of Nigeria Crude Oil Markets using GARCH Model: 1987 – 2017. *International Journal of Applied Science & Mathematical Theory*, 4(3) 23.

Doane, D., Seward, L. (2010) *Applied in Business and Economics*, London ,Mcgraw-Hill, 2010.

Downie, R., (2012). *The state of public health in South Sudan*, S.l: Centre for Strategic and International Studies, Global Health Policy Centre.

Emenike, G. C.,&Ogbole, A.,(2008), "Accidents and the Road Transport Industry in Nigeria,"

Eze. B. (2012) Road Traffic Accidents in Nigeria:A Public Health Problem. *AFRIMEDIC Journal, 3(2)*

Famoye, F. (1993). Restricted generalized Poisson regression model. *Communications in Statistics – Theory and Methods*, 22, 1335-1354.

Ferraro, K. (1996). Women's fear of victimization: shadow of sexual assault? *Social Forces*, 75, 667–90.

Garber, N. J., Wu L. (2001). Stochastic model relating to crash probabilities with geometric and corresponding traffic characteristics data, *UVACTS*-5:15-74

Gardner, W., Mulvey E. P., Shaw E. C. (1995). Regression Analyses of Counts And Rates: Poisson, Over Dispersed Poisson, and Negative Binomial Models. *Psychological Bulletin* 118: 392–404.

Garenne, M., Mbaye, K., Bah, M.D., & Correa, P. (2003). Risk factors for maternal mortality: a case-control study in Dakar Hospitals (Senegal). *African Journal of Reproductive Health*, 1(1), 14-24.

Gonzales, M.M., Dickinson, L.M., DiGuiseppi, L.M., Lowenstein, S.R., (2005), Student drivers: A study of fatal motor vehicle crashes involving 16-year-old drivers. *Ann. Emerg. Med*. 45 (2), 140–146.

Graham, W.J., Fitzmaurice, A.E., Bell, J.S., & Cairns, J.A. (2004). The familial technique for linking maternal death with poverty. *Lancet*, 363(9402), 23-47.

Greene, W (2008). Functional forms for the negative Binomial model for count data. Foundations and Trends in Econometrics. Working Paper, Department of Economics, Stern School of Business, New York University. 585-590.

Greenwood, M. A & Woods H. M. (1919). The incidence of industrial accidents upon special reference to multiple accidents.

Greenwood, M., & Yule, G.U. (1920). An inquiry into the nature of frequency distributions of multiple happenings, with particular reference to the occurrence of multiple attacks of disease or repeated accidents. *Journal of the Royal Statistical Society*, 83, 255-279.

Greenwood. M A & Yule G. U(1920),An inquiry into the Nature Of Frequency Distributions Representative Of Multiple Happenings With Particular Reference To The Occurrence Of Multiple Attacks Of Disease Or Of Repeated Accidents, *Journal Royal Statistics Soc. Ser*. 255-279.

Gupta, P., Gupta, R., &Tripathi, R. (1996). Analysis of zero-adjusted count data. *Computational Statistics & Data Analysis, 23,*207–218. https://doi.org/10.1016/S0167-9473(96)00032-1

Hilbe, J. M. (2011). Negative binomial regression (2nd ed.). Cambridge, UK; New York: Cambridge University Press.

Hinde, J., & Demetrio, C. (1998). Over dispersion: models and estimation. *Computational Statistics & Data Analysis, 27,*151-170.

Hu, M., Pavlicova, M., &Nunes, E. (2011). Zero-inflated and hurdle models of count data with extra zeros: Examples from an HIV-risk reduction intervention trial. *The American Journal of Drug and Alcohol Abuse*, 37(5), 367–375.

Hur, K., Hedeker, D., Henderson, W., Khuri, S., & Daley, J. (2002). Modeling clustered count data with excess zeros in health care outcomes research. *Health Services & Outcomes Research Methodology*, *3*, 5–20.

Hur, K., Hedeker, D., Henderson, W., Khuri, S., & Daley, J. (2002). Modeling clustered count data with excess zeros in health care outcomes research. *Health Services & Outcomes Research Methodology*, *3*, 5–20.

ICDIO, (2016). International statistical classification of disease on health problems, 10th version. Annual Meeting.

Illah, E., Mbaruku, G., Masanja, H. & Kahn, K. (2013). Causes and risk factors for maternal mortality in rural Tanzania - case of Rufiji Health and Demographic Surveillance Site (HDSS). *African Journal of Reproductive Health September,* 17(3), 119-130.

Jahromi, B.N. &Husseini, Z. (2008). Pregnancy outcome at maternal age 40 and older. *Taiwanese Journal of Obstetrics & Gynecology*, 47(3), 318-321.

Jonathan, J.R, Shirley. R,Salissou,&MAidan, F. (2018). What are the factors that contribute to road accidents? An assessment of law enforcement views, ordinary drivers' opinions, and road accident records *Journal of the International Centre for Constructive Research (ICCR).34(27),14-18*

Karen, M. (2015). Time Is Money: An Enquiry Into The Effectiveness Of Road Traffic Management Schemes And Congestion Charges. *Student Economic Review*, 19,( 153)

Kerlinger, F.N (1973). *Foundation of behavioral research*. New York: Holt, Rinehart and Winston.

Kim. S. H., Chung S. B., Song K. S. (2005). Development of an accident prevention model using GLM (Generalized Log -Linear Model) and EB Method a case of seal. *Journal of the Eastern Asia society for Transportation studies* 6:3669-36682.

Lam, L.T., (2003). Factors associated with young drivers' car crash injury: comparisons among learner, provisional, and full licensees. *Accidi Anal. Prev*. 35 (6), 913–920.

Lambert, D. (1992). Zero-inflated Poisson regression with an application to defects in manufacturing. *Technometrics, 34*(1), 1-14.

Lord, D., & Park, P.Y.J. (2008). Investigating the effects of the fixed and varying dispersion parameters of Poisson-Gamma models on empirical Bayes estimates. *Accident Analysis & Prevention*, 40(4), 1441-1457.

Lord, D., Washington, S.P., & Ivan, J.N. (2005). Poisson, Poisson-Gamma and zero inflated regression models of motor vehicle crashes: balancing statistical fit and theory. *Accident Analysis & Prevention,* 37(1), 35-46.

Malyshkina, N., Mannering, F., (2010). Empirical Assessment of the Impact Of Highway Design Exceptions On The Frequency And Severity Of Vehicle Accidents. *Accident Analysis and Prevention* 42(1), 131–139.

Mccartt, A. T., Mayhew, D. R., Braitman, K. A., Ferguson,S. A. & Simpson, H. M.( 2009). Effects of Age and Experience On Young Driver Crashes: Review Of Recent Literature. *Traffic Injury Prevention*,10, 209-219.

McGwin, G., Brown, D.B., (1999). Characteristics of traffic crashes among young, middle aged, and older drivers. *Accid. Anal. Prev.* 31 (3), 181–198.

Miaou, S. P. (1994). The relation between track accident and geometric design of road section: Poisson versus negative binomial regression. Accident Analysis and Prevention

Miaou, S. P., Lum H. (1993). Modeling Vehicle Accidents and Highway Geometric Design Relationships Accident Analysis and Prevention. 25:689-1993

Mojekwu, J.N. (2005). Maternal mortality: natural risk to women. *Ghana Journal of Development Studies*, 2(1), 129-141.

National Population Commission (NPC) [Nigeria] and ICT International (2014). *Nigeria Demographic and Health Survey 2013*. Abuja, Nigeria and Rockville, Maryland, USA: NPC and ICT International.

Newbold E. M (1927), Contribution to the Study of the Human Factor in the Causation of Accidents, Rep. 34, Indust. Fatigue Res. Board, London, 1926.

Nnamdi, M. &Ibekwe, U. (2012). Maternal mortality in Nigeria: examination of intervention methods. *International Journal of Humanities and Social Science*, 2(20), 134-149.

Noland, Robert B., (2001). “Relationships between Highway Capacity and Induced Vehicle Travel”, *Transportation Research A*, 35(1): 47-72.

Nsirim-Worlu, H.G. (2011). Gender stratification: a study of discrimination and oppression in selected communities in Nigeria. *Academic Research International*, 1(2), 238-245.

Nwankwo, C. H &Nwaigwe ,G. I (2016). Statistical model of road traffic crashes data in Anambra State, Nigeria: a Poisson regression approach. European Journal of Statistics and Probability, 1(4), 41-60.

Nwankwo, C. H &Nwaigwe ,G. I (2016). Statistical model of road traffic crashes data in Anambra State, Nigeria: a Poisson regression approach. European Journal of Statistics and Probability, 1(4), 41-60

Nworgu, B.G (2006). *Educational research basic issues and methodology*. Nssuka: University Trust Publishers.

Ogbuokiri, C., &Egeonu, R. (2013). Maternal mortality at NnamdiAzikiwe teaching hospital, southeast Nigeria: a 10-year review (2003 - 2012). *International Journal of Women's Health, 5, 431-436.*

Ohakwe J, Iwueze, I.S. and Chikezie D.C (2011). Analysis of Road Traffic Accidents in Nigeria: A Case Study of Obinze/Nekede/Iheagwa Road in Imo State, Southeastern, Nigeria. *Asian Journal of Applied Sciences, 4: 166-175.*

Olatayo, T. I. &Adeboye, N.O. (2013). Predicting population growth through births and death rate in Nigeria. *Mathematical theory and Modelling*, 3(1), 96-101.

Oluwaseyi J, A &Gbadamosi K, T (2017). Road traffic Crashes in Nigeria: Causes and Consequences. *Transport & Logistics: The International Journal.* 17(42), 2406-1069.

Omoruyi, G. (2008). Causes of maternal mortality in Nigeria, Sunday observer, Benin.

Onwumere, J.U.J. (2005). *Business and economic research methods*. Lagos: Don-Vinton Limited.

Oppong R. A &Assuah C.K(2015), Comparative Assessment Of Poisson And Negative Binomial Regressions As Best Models For Road Count Data *International Journal of Scientific Research and Engineering Studies (IJSRES)*.2 (2349-8862)

Oppong, R. A. (2015). Comparative Assessment of Poisson and Negative Binomial Regressions as Best Models for Road Count Data. *International* Journal of Scientific Research and Engineering Studies, 2(12): 2349-8862.

Oyindamola, B., Yusuf, R., Afolabi, F. &Ayoola, S. A. (2018). Modelling excess zeros in count data with application to antenatal care utilization. *International Journal of Statistics and Probability, 7(3), 62-81*

Philip, N & Sebastian .N,(2017). Application of Poisson Regression on Traffic Safety. Degree Project, in , Second Level Mathematical Statistics stockholm,

Poch, M & Mannering, F. L. (1996). Negative binomial analysis of intersection accident frequency. *Journal of Transportation Engineering*, 122, 105-113.

Ramos, S., Karolinski, A., Romero, M., & Mercer, R. (2007). A comprehensive assessment of maternal deaths in Argentina: translating multicentre collaborative research into action. *Bulletin of the World Health Organization*, 85(8), 615-22.

Road traffic Accident (2013) Retrieved on 4th January,2018fromHttp://en.wikipedia. org/wiki/road_trrafic_accident.

Robert B. N, (2013). From theory to practice in road safety policy: Understanding riskversus mobility, *Research in Transportation Economics* 43, 71-84

Rolison, J. J., Hanoch, Y., & Wood, S. (2014). Risky decision making in younger and older adults: The role of learning. Psychology and Aging, 27, 129–140. Retrieved on 13th April, 2018 from http://dx.doi.org/10.1037/a0024689

Salifu, A. (2014). Analysis of maternal mortality in Wa District. *International Journal of Humanities and Social Science*, 6(3), 34-49.

Sarpong, S. (2012). Analysis of maternal mortality with time: A case study of OkomfoAnokye Teaching Hospital in Kumasi. Master's thesis, Mphil Thesis, Department of Mathematics, Kwame Nkrumah University of Science and Technology

Shankar, V., Mannering, F., Barfield, W.(1995),Effect Of Road Way Geometrics And Environmental Factors On Rural Accident Frequencies. *Accident Analysis and Prevention* 27(3),371–389.

Shankar, V., Milton, J.,& Mannering, F., (1995). Modeling accident frequencies as zero-altered probability processes: An empirical inquiry. Accident Analysis & Prevention 29, 829-837

Sharpie, N. R., De Veaux, R. D., Velleman, P. F. (2010). Business Statistics. Boston, MA: Addison Wesley

Shaw-Pin, M. (1993),The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. Center for Transportation Analysis, Energy Division Oak Ridge National Laboratory

Shiffman, J. &Okonofua, F. (2007). The state of political priority for safe motherhood in Nigeria. *Br. J. Obstet. Gynecol*, 114, 127-133.

Sunday, Punch (2011). Abiye safe delivery, December 25, 2011, 17(19670), 18-19.

Tawiah, K. (2015). Modeling Maternal Mortality in Ghana Using Hierarchical Models (2010-2013).

Ujah, I.A.O., Aisien, O.A., Mutihir, J.T., Vanderagt, D.J., Glew, R.H., &Uguru, V.E. (2005). Factors contributing to maternal mortality in North-Central Nigeria: a seventeen-year review. *African Journal Reproduction Health*, 9(3), 27-40.

Wang, W.R. &Famoye, F. (1997). Modeling household fertility decisions with generalized Poisson regression. *Journal of Population Economics*,10(3), 273-283

WHO (2004). Factsheet on Road Traffic Injuries from https//www.who.int

WHO (2015). Fact sheets on sustainable development goals : health targets. Maternal health. Retrieved from http://www.euro.who.int/__data/assets/pdf_ file/0006/354921/3.1-SDG-Fact-sheet-Maternal-Health.pdf?ua=1

WHO, (2016). Countries: South Sudan. [Online] Availableat:http://www.who.int/ countries/ssd/en/ [Accessed 8 January 2017].

Wimberley, D.W. & Bello, R. (1992). Effects of foreign investment, exports, and economic growth on third world food consumption. *Social Forces*,70, 895-921.

Wodi, B.E. (2005). HIV/AIDS knowledge, attitudes, and opinions among adolescents in the River States of Nigeria. *The International Electronic Journal of Health Education*, 8, 86-94.

World Health Organization (2008), Factsheet, maternal mortality, Department of Making Pregnancy Safer.

World Health Organization (2013). WHO global database on maternal health indicators, 2013 update. Geneva, World Health Organization. Retrieved from http://www.who.int/gho.

Xiaoxiang, M, (2016). Refined-Scale Crash Data Analysis Using Multi-Level Regression Models. Dissertation Department of Civil And Environmental Engineering In Partial Fulfillment Of The Requirements For The Degree of Doctor of Philosophy Colorado State University Fort Collins, Colorado

Xuedong Y, B, W, Meiwu, A, &CuipingZ(2012), distinguishing between Rural andUrban Road Segment Traffic Safety Based onZero-Inflated Negative BinomialRegression

Models.China, Beijing Hindawi Publishing Corporation Discrete Dynamics in Nature and Society. 2(11)

Yusuf, O. B., &Ughali, L. (2018). On the performance of the Poisson, negative binomial and generalized Poisson regression models in the prediction of antenatal care visits in Nigeria. *American Journal of Mathematics and Statistics, 5*(3), 128-136.