Modelling Excess Zeros in Count Data with Application to Antenatal Care Utilisation

Oyindamola B. Yusuf[1]. Rotimi F. Afolabi[1] & Ayoola S. Agbaje[1]

[1] Department of Epidemiology and Medical Statistics, Faculty of Public Health, College of Medicine, University of Ibadan, Nigeria

Correspondence: Oyindamola B. Yusuf, Department of Epidemiology and Medical Statistics, Faculty of Public Health, College of Medicine, University of Ibadan, Nigeria.

Abstract

Poisson and negative binomial regression models have been used as a standard for modelling count outcomes; but these methods do not take into account the problems associated with excess zeros. However, zero-inflated and hurdle models have been proposed to model count data with excess zeros. The study therefore compared the performance of Zero-inflated (Zero-inflated Poisson (ZIP) and Zero-inflated negative binomial (ZINB)), and hurdle (Hurdle Poisson (HP) and Hurdle negative binomial (HNB)) models in determining the factors associated with the number of Antenatal Care (ANC) visits in Nigeria. Using the 2013 Nigeria Demographic and Health Survey dataset, a sample of 19 652 women of reproductive age who gave birth five years prior to the survey and provided information about ANC visits was utilised. Data were analysed using descriptive statistics, ZIP, ZINB, HP and HNB models, and information criteria (AIC/BIC) was used to assess model fit. Participants' mean age was 29.5 ± 7.3 years and median number of ANC visits was 4 (range: 0 - 30). About half (54.9%) of the participants had at least 4 ANC visits while 33.9% had none. The ZINB (AIC = 83 039.4; BIC = 83 470.3) fitted the data better than the ZIP or HP; however, HNB (AIC = 83 041.4; BIC = 83 472.3) competed favorably well with it. The Zero-inflated negative binomial model provided the better fit for the data. We suggest the Zero-inflated negative binomial model for count data with excess zeros of unknown sources such as the number of ANC visits in Nigeria.

Keywords: antenatal care utilisation, zero-inflated Poisson, zero-inflated negative binomial, hurdle Poisson, hurdle negative binomial, excess zeros

1. Introduction

Poisson regression and Negative Binomial regression models have been used as a standard for modelling count outcomes; but these methods do not take into account the problems associated with over dispersion and excess zeros (Hinde & Demetrio, 1998). When excess zeros are observed in count data, studies have suggested that applying the customary regression model is a problem (Hinde & Demetrio, 1998; Hur et al., 2002). Previous studies (Gupta, Gupta, & Tripathi, 1996; Hur et al., 2002) have shown that more error is observed for small values of the count as a result of failing to account for excess zeros and this may result in biased parameter estimates and wrong inferences.

In addressing these problems, diverse modifications of basic regression models have been proposed. Notable among the modifications are hurdle and Zero-inflated regression models (Poisson or negative binomial distribution of the outcome assumed in either case) that possess a lot of potentials including modelling excess zeros (Hu, Pavlicova, & Nunes, 2011; Staub & Winkelmann, 2013). As reported by Xu, et al. (2015), the hurdle and Zero-inflated count models show unbiased estimations. Hurdle models – Hurdle Poisson (HP) or Hurdle Negative Binomial (HNB) - determine whether the outcome is zero or nonzero (Mullahy, 1986). The model assumes only one source of zero (structural zero). On the other hand, Zero-inflated models - Zero-inflated Poisson (ZIP) or Zero-inflated Negative Binomial (ZINB) – are based on the premises that the data come from two distinct populations: one population that never experiences the event, and another one for which events are generated from a standard model (Lambert, 1992). Unlike hurdle models, zero-inflated models identify two sources of zeros (structural and sampling zeros) which may not be generally distinguished. These regression models have been applied as tools for modelling data with excess zeros in recent years. Meanwhile, the number of Antenatal Care (ANC) visits is an example of such count data with excess zeros.

By and large, ANC utilisation serves as an indicator for assessing the proportion of pregnant women who received ANC from a healthcare provider at least once and/or at least four times recommended during pregnancy (WHO, 2013). Globally,

it is a major factor that contributes to the burden of maternal mortality. Meanwhile, UNICEF (2016) reported that about 85% of pregnant women accessed ANC at least once while only 58% received at least four ANC visits with a skilled healthcare provider as recommended by the WHO. Low rates of maternal mortality reported for developed countries compared with the high figures in developing countries have been attributed to the higher utilisation of ANC by the developed countries (Cook et al., 2001; WHO, 2105). Meanwhile, as reported by Sedgh et al. (2014), out of about 213 million pregnancies which occurred worldwide in 2012, 89% (190 million) occurred in the developing world. In recent years, ANC utilisation is about 4% in sub-Saharan Africa, and there has been slower progress in improvements compared with other regions (Lincetto et al., 2010). In South Sudan, about two-fifths of the women utilised ANC (Mugo, Dibley & Agho, 2015) while majority of women did not receive optimum number of ANC visits in Ethiopia (Kawungezi et al., 2015). Less than half (49%) of the pregnant women in sub-Saharan Africa utilised ANC facility despite the fact that WHO recommends a minimum of four ANC visits (Kawungezi et al., 2015). In Nigeria, maternal mortality has been described as a vital indicator for women's programmes as well as reproductive health programmes (NPC & ICT, 2014). As reported in 2013 NDHS, the burden of maternal mortality is 576 maternal deaths per 100,000 live births in Nigeria. Nigeria makes up 1.7% of the total world population yet, contributes 10% of global maternal mortality and ranks second globally (to India) in the number of maternal deaths. This is worrisome in that every 1 in 30 women die as a result of pregnancy or childbearing related deaths (NPC & ICT, 2014). Antenatal care was not only described as a major determinant of high maternal mortality rate but also as one of the basic components of maternal care (Nisar & White, 2003). Proper utilisation of ANC thus enhances safe motherhood and delivery with improved maternal outcomes.

Most of the previous studies on ANC utilisation in Nigeria have examined the factors affecting ANC utilisation using the binary logistic regression techniques by treating ANC utilisation as a categorical variable (yes/no) (Adamu, 2011; Awusi, Anyanwu, & Okeleke., 2009; Dairo & Owoyokun, 2010; Onasoga, Afolayan, & Oladimeji, 2012). On the other hand, Yusuf & Ugalahi (2015) identified factors affecting ANC utilisation using a model for handling over-dispersion without taking into consideration excess zeros in the data. Although, there is a likelihood that the presence of excess zeros in count data suggests over-dispersion, Greene (1994) suggested distinct models for zero-inflated and over-dispersed datasets. Hence, the present study was designed to explore the factors affecting ANC utilisation in Nigeria using the Zero-inflated and Hurdle regression models.

2. Methods

2.1 Data Extraction

Data for this study was extracted from the 2013 Nigeria Demographic and Health Survey (NDHS) dataset. The survey was a cross-sectional population-based study. The present study population were women (n=19652) of childbearing age (15-49 years) who gave birth five years prior to the survey and provided information about ANC visits. Sampling and data collection methods have been previously reported (NPC & ICT, 2014).

2.2 Statistical Analysis

Descriptive statistics were used to summarize the outcome variable (number of ANC visits) and socio-demographic characteristics of the participants. Both the Zero-inflated regression and Hurdle models' parameters were estimated. The explanatory variables used were the socio-demographic characteristics such as age, residence, wealth quintile, etc. Test of dispersion parameter was conducted using Wald test of ln(alpha) to compare nested models: ZIP against ZINB; and HP against HNB. Non-nested models ZIP/HP against Poisson and ZINB/HNB against Negative binomial were also compared by applying Vuoug test to ascertain the existence of excess zeros (Vuong, 1989). Model fits were thereafter assessed using both the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). All analyses were performed at 5% significance level using the Stata software (Stata, 2011).

2.3 Statistical Modelling

Zero-inflated (ZIP or ZINB) and Hurdle (HP or HNB) models have been developed to deal with the excess zeros in the observed outcome data; however, they are distinctively different in the interpretation and analysis of such count data (Hu et al, 2011). Both models combine binomial probabilities with either Poisson or negative binomial distributions. These models are sometimes referred to as 2-part models (data generating processes). Considering ANC utilisation measured by the number of ANC visits, a pregnant woman might have had zero ANC visits during a given period probably because of the two data generating processes.

2.3.1 Zero-inflated Regression Model

Let $Y_i, i = 1, 2, \cdots, n$, be the number of ANC visits (the outcome variable). For every observation i, the first data generating process generates only zero counts (structural zeros), with probability φ_i while the second generates non-negative counts, which could result in zero counts (sampling zeros), from either a Poisson or a negative binomial model with probability $(1 - \varphi_i)$. In this present study, specifically, the two possible processes are that a woman who is

pregnant will attend ANC or that she will not. If a pregnant woman does not attend ANC then the only possible outcome is zero. She either never visits (structural zeros) or she could visit but has no reason to do so within the specified time frame (sampling zeros). On the other hand, if a pregnant woman attends ANC; it is a count (nonzero) process. These two parts of a zero-inflated model are regarded as a binary model (usually a logit model) and a count model (in this case, a Poisson model or negative binomial model). In the light of this, the distribution of number of ANC visits can be modelled effectively using Zero-inflated regression model with Y_i defined as

$$y_i \sim \begin{cases} 0 \text{ (structural - with probabilit y } \varphi_i\text{; or } sampling - \text{with probabilit y } 1\text{-}\varphi_i) \\ \geq 1 \text{ (nonzero count - with probabilit y } 1\text{-}\varphi_i) \end{cases} \tag{1}$$

where φ_i (zero inflation parameter) is the probability of zero (structural) counts from the binary process; in any dataset, whenever $\varphi_i > 0$ excess zeros or zero-inflation occurs (Staub & Winkelmann, 2013).

The above equation (1) can be referred to as an indicator function in that if the observed response $y_i = 0$ it takes the value of 1; while a value of 0 if $y_i \geq 1$. With the notion that the overall probability of zero counts is the combined probability of zeros from the two processes defined above, probability of $Y_i = y_i$ can therefore be described as:

$$P(Y_i = y_i) = \varphi_i + (1-\varphi_i)g(y_i), y_i = 0 \tag{2}$$

$$P(Y_i = y_i) = (1-\varphi_i)g(y_i), y_i \geq 1 \tag{3}$$

where $g(y_i)$ is the probability function of either the Poisson or the negative binomial. For $E_h(y_i)$ denoting the mean of the parent distribution, the expected value of the zero-inflated count model is expressed as:

$$E(y_i) = (1-\varphi_i)E_h(y_i) \tag{4}$$

Specifically, zero-inflated regression models allow for covariates for both the binary and count parts of the model owing to the fact that the covariates that influence φ_i and $g(y_i)$ parameter are not necessarily equivalent.

Suppose β and γ are the vectors of unknown parameters associated with the known covariate vectors $Z_i = \begin{pmatrix} 1 & z_{i1} & \cdots & z_{iq} \end{pmatrix}^{'}$ and $X_i = \begin{pmatrix} 1 & x_{i1} & \cdots & x_{ip} \end{pmatrix}^{'}$ respectively. Traditionally, φ_i is expressed in terms of a logistic regression model which depends on the characteristics of observation i as

$$\log it(\varphi_i) = Z_i^{'}\gamma \tag{5}$$

where Z_i is a $1 \times (q+1)$ vector of zero-inflated covariates for the logistic component, γ is the corresponding $(q+1) \times 1$ vector of logistic regression coefficients, and q is the number of covariates.

$$\log(\mu_i) = X_i^{'}\beta \Rightarrow \mu_i = \exp(X_i^{'}\beta) \tag{6}$$

where X_i is a $1 \times (p+1)$ vector of zero-inflated covariates for the parent distribution component, β is the corresponding $(p+1) \times 1$ vector of regression coefficients, and p is the number of covariates.

Let the function F (a scalar) relating the product $Z_i^{'}\gamma$ to the probability φ_i be the zero-inflated link function such that,

$$\varphi_i = F_i = F(Z_i^{'}\gamma) \tag{7}$$

Thus, specifying the zero-inflated link function F as the logistic function, equation (7) becomes

$$F(Z_i^{'}\gamma) = \frac{\exp(Z_i^{'}\gamma)}{1 + \exp(Z_i^{'}\gamma)} \tag{8}$$

Essentially, zero-inflated models allows researchers to find answers to these two basic questions regarding the number of ANC visits with excess zeros: (i) what predicts whether or not a pregnant woman will attend ANC, and (ii) what predicts the number of times she may visit ANC? Zero-inflated models therefore have more statistical potentials compared with the standard regression models in that they model both the preponderance of zeros and the distribution of counts at the same time (Lord, 2006).

Zero-inflated Poisson model: A ZIP model, introduced by Lambert (1992), allows for covariates for both the binary and Poisson parts of the model, and it has been commonly used to model count data with excess zeros (Hur et al., 2002). It assumes that with probability φ_i the only possible observation is zero, and with probability $(1-\varphi_i)$ a Poisson (λ) distribution is observed. The other data (count) generating process, for a potential pregnant woman who could visit ANC, has a Poisson distribution with parameter λ=μ, and its probability mass function expressed as

$$g(y_i) = \frac{\exp(-\mu_i)\mu_i^{y_i}}{y_i!} \tag{9}$$

It follows that the number of events y has a Poisson distribution with a conditional mean that depends on an individual's characteristics, x_i such that

$$E_h(y_i / x_i) = \mu_i = \lambda_i \tag{10}$$

where $\log(\lambda_i) = X_i'\beta \Leftrightarrow \lambda_i = \exp(X_i'\beta)$

With the probability mass function of count process, as defined above in equation (9), the ZIP model is expressed as

$$P(y_i / x_i, z_i) = \begin{cases} F_i + (1 - F_i)\exp(-\mu_i), y_i = 0 \\ (1 - F_i)\dfrac{\exp(\mu_i)\mu_i^{y_i}}{y_i!}, y_i \geq 1 \end{cases} \tag{11}$$

The conditional mean of y_i and the conditional variance are respectively defined as follows.

$$E(y_i / x_i, z_i) = \mu_i(1 - F_I) \tag{12}$$

$$V(y_i / x_i, z_i) = E(y_i / x_i, z_i)(1 + \mu_i F_I) \tag{13}$$

Zero-inflated negative binomial model: ZINB models have been described as an extended version of the negative binomial regression models for excess zero count data (Greene, 1994). Considering the replacement of equation (9) with the negative binomial regression model, according to Greene, the ZINB model can be derived. Thus, by specifying a negative binomial distribution for the second data (count) generation process (as described above),

$$g(y_i) = \frac{\Gamma(\dfrac{1 + \alpha y_i}{\alpha})}{\Gamma(\dfrac{1}{\alpha})y_i!}(\frac{1}{1 + \alpha\mu_i})^{\frac{1}{\alpha}}(\frac{\alpha\mu_i}{1 + \alpha\mu_i})^{y_i} \tag{14}$$

The ZINB model is thus expressed as

$$P(y_i / x_i, z_i) = \begin{cases} F_i + (1 - F_i)(\dfrac{1}{1 + \alpha\mu_i})^{\frac{1}{\alpha}}, y_i = 0 \\ (1 - F_i)\dfrac{\Gamma(\dfrac{1 + \alpha y_i}{\alpha})}{\Gamma(\dfrac{1}{\alpha})y_i!}(\frac{1}{1 + \alpha\mu_i})^{\frac{1}{\alpha}}(\frac{\alpha\mu_i}{1 + \alpha\mu_i})^{y_i}, y_i \geq 1 \end{cases} \tag{15}$$

where α (alpha) is the over-dispersion parameter since increasing α increases the conditional variance of y, and both the conditional mean of y_i and conditional variance are expressed respectively as follows,

$$E(y_i / x_i, z_i) = \mu_i(1 - F_I) \tag{16}$$
$$V(y_i / x_i, z_i) = E(y_i / x_i, z_i)[1 + \mu_i(F_I + \alpha)] \tag{17}$$

Note that, if α = 0, the mean and variance are the same, and we have a Poisson model. As with the ZIP model, the ZINB model corrects for over dispersion that arises when the variance is greater than the conditional mean.

2.3.2 Hurdle Models

A hurdle model (HP/HNB) is a 2-part model of which the first part models whether a zero or non-zero response outcome occurs based on a binomial model; while the second part, conditioning on it being positive count outcomes, models the truncated count distribution (either Poisson or negative binomial). Essentially, hurdle models simultaneously fit two separate regression models (Dalrymple et al., 2003; He et al., 2014).

Suppose a pregnant woman had zero ANC visit during a given period because she never visits; the two disjoint modelling stages are as follows. One, a binary model part which determines whether a woman visits ANC (positive count outcome) or not (structural zero outcome) and logistic regression is usually used to allow for the investigation of the effects (βi) of covariates on the probability of an observation being zero (π_i). Two, it is a stage that models positive (nonzero) count truncated at-zero count outcomes following either truncated Poisson or negative binomial distribution.

Similarly to the zero-inflated model, the distribution of number of ANC visits can be modelled effectively using hurdle model with Y_i defined as

$$y_i \sim \begin{cases} 0 \text{ (structural - with probabilit y } \pi_i) \\ \geq 1 \text{ (nonzero count - with probabilit y } 1 - \pi_i) \end{cases} \tag{18}$$

where π_i (zero inflation parameter) is the probability of zeros.

Considering $Y_i, i = 1, 2, \cdots, n$, as the number of ANC visits (the outcome variable) of which the conditional distribution is controlled by standard truncated-at-zero distribution, the probability of $Y_i = y_i$ can therefore be described as (Mullahy, 1986):

$$P(Y_i = y_i) = \pi_i, y_i = 0 \tag{19}$$

$$P(Y_i = y_i) = (1 - \pi_i) \frac{g(y_i)}{1 - g(0)}, y_i \geq 1 \tag{20}$$

where $g(y_i)$ denotes the probability function of a standard count data model, usually either the Poisson or the negative binomial distribution (which models the frequency of repeat visits for y≥1).

Of note, formulation of hurdle (HP/HNB) model is similar to the ZIP/ZINB model: the models combine binomial probabilities with Poisson/negative binomial distributions. Essentially, HP (HNB) can be considered as a reparameterisation of ZIP (ZINB) if φ_i is constant across the samples. However, unlike ZIP/ZINB model, HP/HNB model separates the zeros from the non-zeros response by modelling the $y_i \geq 1$ with a truncated Poisson/negative binomial distribution. Besides, as argued by Hu et al (2011), the effects of covariates on π_i in a hurdle model (on the log-odds of a zero response) and their effects on φ_i in a ZIP model (on the log-odds of structural zeros) are not equivalent.

2.3.3 Model Selection Technique

Maximum likelihood model-based methods have been suggested in the literature to compare the performance of alternative models (Rideout et al., 2001; Staub and Winkelmann, 2013). For instance, information criterion measures supply information on the strength of evidence for each model; it does not make use of a significant level rather it is based on maximum likelihood. In this study, commonly used information criterion-based measures – AIC and BIC were used to assess models' performances. The main objective of the measures is to identify the model that most reasonably generated the data, and it is mathematically expressed as:

$$AIC = -2Log(L) + 2p \tag{21}$$
$$BIC = -2Log(L) + \log(n)p \tag{22}$$

where p = number of estimated parameters included in the model, n= the total number of observations and L = likelihood function

In summary, either of the aforementioned information criteria shows how well the model fits the data or set of observations. The smaller the AIC or BIC, the better is the model.

3. Results

3.1 Number of ANC Visits

Of the 19 652 women of child bearing age from whom valid and complete information on ANC visits were obtained, median number of ANC visits was 4 and range, 0 – 30 visits. The variance (37.237) was greater than the mean (5.26) number of ANC visits, and number of zeros present in the data was 6 662 (33.9%) as shown in Figure 1.
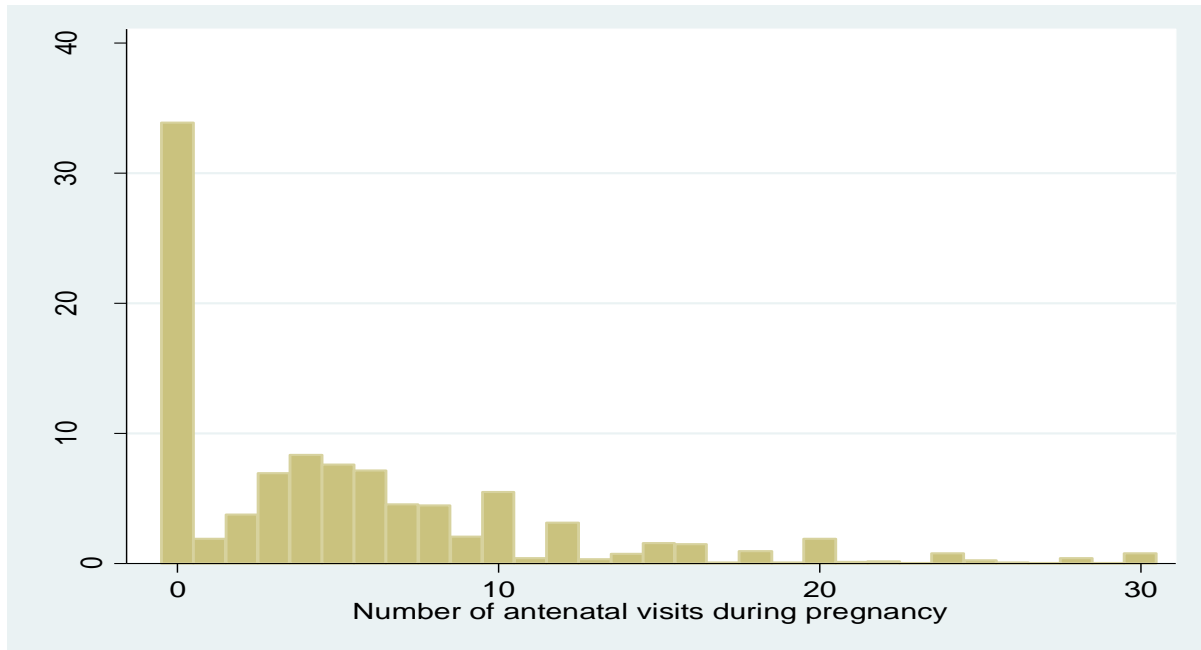

Figure 1. Distribution of the number of ANC visits by the participants

3.2 Socio-demographic Characteristics of the Respondents

Table 1 shows that the mean age of the total respondents (29.5±7.3 years) was nearly equivalent to those who had zero ANC visit (29.9±7.7 years). By age group, the highest proportion of women aged 15-24 years (38.2%) and those aged 40-49 years (37.3%) had zero ANC visit. Most women from the North West (58.4%) and North East (37.1%) also had zero ANC visit compared with those from South East (5.0%) and South West (7.1%). About 45% of respondents who resided in rural areas compared with 11.2% in urban residence had zero ANC visit. In addition, more than half (57.0%) of those who had no education compared with women of higher education (1.6%) had zero ANC visit. About two-thirds (67.8%) of women in the poorest wealth quintile compared with a few (3.9%) in richest wealth quintile had zero ANC visit. Most unemployed participants (43.5%) compared with 29.6% of those employed did not utilise ANC. Unlike the respondent's employment status, most women whose husbands/partners had employment (34.4%) had zero ANC visit compared to 23.8% of those who were not employed. Proportion of women with zero ANC visit reduces as husbands/partners education level advanced: no education (61.1%), primary education (26.9%), secondary (16.6%) and higher education (7.9%).

Table 1. Socio-demographic characteristics of the respondents by non utilisation of ANC

Variables	Zero ANC Visit (%)	Total (%)
Age group(years)		
15 – 24	1 935 (38.2)	5 068 (25.8)
25 – 39	3 866 (31.5)	12 273 (62.4)
40 – 49	861 (37.3)	2 311(11.8)
Mean ± sd	29.9 ±7.7	29.5 ±7.3
Region		
North Central	656 (21.7)	3 025 (15.4)
North East	1 454 (37.1)	3 917 (19.9)
North West	3 590 (58.4)	6 145 (31.3)
South East	82 (5.0)	1 653 (8.4)
South South	698 (29.8)	2 343 (11.9)
South West	182 (7.1)	2 569 (13.1)
Residence		
Urban	726 (11.2)	6 503 (33.1)
Rural	5 936 (45.1)	13 149 (66.9)
Educational level		
No education	5 156 (57.0)	9 042 (46.0)
Primary	917 (23.0)	3 992 (20.3)
Secondary	569 (10.7)	5 331 (27.1)
Higher	20 (1.6)	1 287 (6.6)
Religion		
Christian	1 373 (17.2)	8 005 (40.9)
Islam	5 137 (45.3)	11 346 (58.0)
Traditional	117 (55.7)	210 (1.1)
Wealth index		
Poorest	2 937 (67.8)	4 329 (22.0)
Poorer	2 120 (46.9)	4 522 (23.0)
Middle	1 010 (25.4)	3 975 (20.3)
Richer	471 (12.9)	3 662 (18.6)
Richest	124 (3.9)	3 164 (16.1)
Parity		
1	958 (27.1)	3 541 (18.0)
2 – 4	2 651 (30.7)	8 622 (43.9)
>4	717 (34.0)	2 111 (38.1)
Mean ± sd	4.7±2.9	4.1±2.7
Wanted last child*		
Yes	6 188 (35.4)	17 502 (89.5)
No	410 (19.9)	2 061 (10.5)
Current marital status		
Married	6 248 (34.8)	17 941 (91.3)
Not married	414 (24.2)	1 711 (8.7)
Respondent's employment status*		
Unemployed	2 644 (43.5)	6 074 (31.0)
Employed	3 988 (29.6)	13 493 (69.0)
Husband/Partner's educational level*		
No education	4 386 (61.1)	7 184 (38.2)
Primary	967 (26.9)	3 600 (18.9)
Secondary	928 (16.6)	5 595 (29.4)
Higher	203 (7.9)	2 572 (13.5)
Husband/Partner's employment status*		
Unemployed	43 (23.8)	181 (1.0)
Employed	6 476 (34.4)	18 833 (99.0)
Total	6 662 (33.9)	19 652

  *missing value not reported   sd = standard deviation

3.3 Model Selection of ANC Utilisation

As shown by the Wald's (Z) statistic (p < 0.001) in Table 2, standard negative binomial-based models were significantly favoured against their standard Poisson-based counterparts owing to the evidence of over-dispersion. Comparing excess zeros-based models - ZIP/HP with the standard Poisson model (ZIP Vuong: Z = 58.07; p < 0.001) and ZINB/HNB with negative binomial model (ZINB Vuong: Z = 66.90; p < 0.001), both the hurdle and the zero-inflated models were significantly fit. Of the four different models (ZIP, ZINB, HP and HNB) fitted, ZINB had the least AIC (83 039.4) and BIC (83 470.32). However, HNB (AIC = 83 041.4 and BIC = 83 472.32) competed favorably well with ZINB.

Table 2. Test and fit statistics

Model	Log(L)	AIC	BIC	Vuong (Z)	Wald (Z)
ZIP	-44 846.2	89 800.40	90 223.48	58.07*	
ZINB	-41 464.7	**83 039.40**	**83 470.32**	62.90 *	-73.74 *
HP	-44 846.0	89 799.94	90 223.03		
HNB	-41 465.7	83 041.40	83 472.32		-73.66*

*significant at p<0.001

3.4 Factors Associated with ANC Utilisation

The parameter estimates of some of the factors affecting number of ANC visits using ZINB and HNB models are presented in Table 3. The respective parameter estimates of these models are rather similar except in their interpretation. However, ZIP and HP parameter estimates are not presented in this study.

3.4.1 Application of ZINB Model

Count group: Women aged 25-39 years (IRR = 1.056; 95% CI: 1.03, 1.09) and those aged 40-49years (IRR = 1.095; 95% CI: 1.05, 1.15) utilised ANC more compared to women aged less than 24 years. Women from the South South (IRR = 1.157; 95% CI: 1.11, 1.20), South East (IRR = 1.309; 95% CI: 1.26, 1.36), and South West (IRR = 1.891; 95% CI: 1.83, 1.96) zones respectively had a 15.7%, 30.9% and 89.1% increased rate of ANC visits; while those from the North East (IRR = 0.748; 95% CI: 0.72, 0.78) and North West (IRR = 0.731; 95% CI: 0.70, 0.76) zones respectively had a 25.2% and 26.9% reduction compared to women from the North Central zone. Women who had primary (IRR = 1.037; 95% CI: 1.00, 1.07), secondary (IRR = 1.109; 95% CI: 1.07, 1.15) and higher (IRR = 1.161; 95% CI: 1.10, 1.22) education respectively had a 3.7%, 10.9% and 16.1% increased in ANC visits compared with women having no education. Likewise, women in the middle (IRR = 1.117; 95% CI: 1.07, 1.17), richer (IRR = 1.204; 95% CI: 1.15, 1.26) and richest (IRR = 1.254; 95% CI: 1.19, 1.32) wealth quintile had increased rate in ANC utilisation compared to women in the poorest category. On the other hand, women who had "2 – 4" (IRR = 0.946; 95% CI: 0.92, 0.97) and above 4 (IRR = 0.946; 95% CI: 0.91, 0.98) total children ever born utilised ANC less often than women who ever had only one child. Meanwhile, women whose husband/partner's had primary (IRR = 1.042; 95% CI: 1.01, 1.08), secondary (IRR = 1.083; 95% CI: 1.04, 1.12) and higher (IRR = 1.096; 95% CI: 1.05, 1.14) education respectively had a 4.2%, 8.3% and 9.6% increase in ANC visits compared with no education. Nearly all the independent variables considered had statistically significant effect on the number of ANC visits; however, being a rural dweller (IRR = 0.995, p = 0.712), unmarried (IRR = 1.014, p = 0.478) or having a husband/partner who is employed (IRR = 0.937, p = 0.165) did not significantly affect the number of ANC visits.

Inflation zero group: Women from the South East (e$^{-1.043}$ (OR) = 0.352; p < 0.001), North East (OR = 0.668; p < 0.001), and South West (OR = 0.691; p = 0.001) zones had reduced odds of not visiting ANC compared with those from the North Central zone. However, women from the North West (OR = 1.863; p < 0.001) and South South (OR = 4.039; p < 0.001) zones had increased odds of not attending ANC. Compared to urban women, being a rural dweller (OR = 1.733; p < 0.001) increases the odds of ANC non-utilisation by about 2 times. Women who had primary (OR = 0.487; p < 0.001), secondary (OR = 0.299; p < 0.001) and higher (OR = 0.077; p < 0.001) education had decreased odds of not attending ANC compared with those of no education category. Again, women in the poorer (OR = 0.545; p < 0.001), middle (OR = 0.302; p < 0.001), richer (OR = 0.200; p < 0.001) and richest (OR = 0.112; p < 0.001) wealth quintiles respectively had a 45.5%, 69.8.0%, 80.0 % and 88.8 % reduction in the odds of not attending ANC compared with those in the poorest category. On the other hand, being employed (OR = 0.689; p < 0.001) reduced the odds of ANC non-utilisation by about 31.1% compared to an unemployed woman. Similarly, women whose husband/partner's had primary (OR = 0.490, p < 0.001), secondary (OR = 0.516, p < 0.001) and higher (IRR = 0.387, 95% CI: 0.32, 0.47) education respectively had a 51.0%, 48.4% and 61.3% reduction in the odds of not utilising ANC compared to those whose husbands/partners had no education.

3.4.2 Application of HNB model

Count group: Women aged 25-39 years (IRR = 1.057; 95% CI: 1.03, 1.09) and those aged 40-49 years (IRR = 1.096; 95% CI: 1.05, 1.15) utilised ANC more compared to women aged less than 24 years. Women from the South South (IRR = 1.156; 95% CI: 1.11, 1.20), South East (IRR = 1.308; 95% CI: 1.26, 1.36) and South West (IRR = 1.888; 95% CI: 1.83, 1.95) had increased rate of ANC visits; those from the North East (IRR = 0.745; 95% CI: 0.72, 0.77) and North West (IRR = 0.730; 95% CI: 0.70, 0.76) had a reduction compared with those from the North Central zone. Women who had primary (IRR = 1.037; 95% CI: 1.00, 1.07), secondary (IRR = 1.109; 95% CI: 1.07, 1.15) and higher (IRR = 1.163; 95% CI: 1.11, 1.22) education respectively had a 3.7%, 10.9% and 16.3% increased rate in ANC utilisation compared with women having no education. Likewise, women in the middle (IRR = 1.119; 95% CI: 1.07, 1.17), richer (IRR = 1.204; 95% CI: 1.15, 1.26) and richest (IRR = 1.254; 95% CI: 1.19, 1.32) wealth quintile had increased rate in ANC utilisation compared to women in the poorest category. On the other hand, women who had "2 – 4" and above 4 children ever born utilised ANC less often (IRR = 0.946; 95% CI: 0.91, 0.98) than those who had only one child. Meanwhile, women whose husband/partner's had primary (IRR = 1.042; 95% CI: 1.00, 1.08), secondary (IRR = 1.083; 95% CI: 1.04, 1.12) and higher (IRR = 1.096; 95% CI: 1.05, 1.14) education respectively had a 4.2%, 8.3% and 9.6% increased rate of ANC visits compared with those with no education.

Inflation zero group: Women aged 25-39 years ($e^{-0.115}$ (OR) = 0.891; p = 0.042) and those aged 40-49 years ($e^{-0.179}$ (OR)= 0.836; p = 0.030) had reduced odds of not utilising ANC compared to those aged less than 24 years. Women from the South East (OR = 0.352; p < 0.001), North East (OR = 0.719; p < 0.001), and South West (OR = 0.629; p < 0.001) had reduced odds of ANC non-utilisation. Those from the North West (OR = 1.919; p < 0.001) and South South (OR = 3.645; p < 0.001) had about twice and four times odds of not attending ANC compared to those from the North Central zone. Compared to urban women, being a rural dweller (OR = 1.655; p < 0.001) increased the odds of ANC non-utilisation by about 2 times. Women who had primary (OR = 0.505; p < 0.001), secondary (OR = 0.317; ; p < 0.001) and higher (OR = 0.115; p < 0.001) education had decreased odds of not attending ANC compared with those of no education category. Again, women in the poorer (OR = 0.557; p < 0.001), middle (OR = 0.312; p < 0.001), richer (OR = 0.206; p < 0.001) and richest (OR = 0.121; p < 0.001) wealth quintiles respectively had a 44.3%, 68.8.0%, 79.4 % and 87.9 % reduction in the odds of not attending ANC compared with those in the poorest category. On the other hand, being employed (OR = 0.700; p < 0.001) reduced the odds of ANC non-utilisation by about 30 % compared to an unemployed woman. Similarly, women whose husband/partner's had primary (OR = 0.503; p < 0.001), secondary (OR = 0.521; p < 0.001) and higher (OR = 0.401; p < 0.001) education respectively had a 49.7%, 47.9% and 59.9% reduction in the odds of not attending ANC compared to those whose husband/partner had no education.

Table 3. Parameter estimates of some factors affecting ANC using excess zero-based models: Hurdle negative binomial regression and Zero-inflated negative binomial regression models

Characteristics	HNB model					ZINB model				
Count group	IRR	SE	P-value	95%	CI for IRR	IRR	SE	P-value	95%	CI for IRR
Constant	5.544	0.310	0.000	4.968	6.186	5.505	0.308	0.000	4.934	6.142
Age group										
15 – 24*										
25 - 39	1.057	0.016	0.000	1.025	1.089	1.056	0.016	0.000	1.025	1.089
40 - 49	1.096	0.025	0.000	1.048	1.146	1.095	0.025	0.000	1.047	1.145
Region										
North Central*										
North East	0.745	0.014	0.000	0.718	0.774	0.748	0.014	0.000	0.721	0.777
North West	0.730	0.014	0.000	0.703	0.758	0.731	0.014	0.000	0.704	0.759
South East	1.308	0.027	0.000	1.256	1.362	1.309	0.027	0.000	1.258	1.363
South South	1.156	0.024	0.000	1.110	1.203	1.157	0.024	0.000	1.111	1.204
South West	1.888	0.032	0.000	1.827	1.952	1.891	0.032	0.000	1.830	1.955
Residence										
Urban*										
Rural	0.993	0.013	0.594	0.969	1.018	0.995	0.013	0.712	0.971	1.020
Educational level										
No education*										
Primary	1.037	0.018	0.039	1.002	1.073	1.037	0.018	0.035	1.003	1.073
Secondary	1.109	0.021	0.000	1.069	1.151	1.109	0.021	0.000	1.069	1.150
Higher	1.163	0.030	0.000	1.105	1.224	1.161	0.030	0.000	1.103	1.221
Religion										
Christian*										

Islam	1.047	0.015	0.002	1.017	1.076	1.046	0.015	0.002	1.017	1.075
Traditional/Other	0.986	0.059	0.816	0.876	1.110	0.985	0.060	0.802	0.875	1.109
Wealth index										
Poorest*										
Poorer	1.020	0.023	0.389	0.975	1.066	1.018	0.023	0.436	0.974	1.064
Middle	1.119	0.026	0.000	1.069	1.171	1.117	0.026	0.000	1.067	1.169
Richer	1.204	0.030	0.000	1.147	1.263	1.204	0.030	0.000	1.147	1.264
Richest	1.254	0.034	0.000	1.188	1.323	1.254	0.034	0.000	1.188	1.323
Parity										
1*										
2 - 4	0.946	0.014	0.000	0.918	0.975	0.946	0.014	0.000	0.918	0.974
More than 4	0.946	0.018	0.003	0.912	0.982	0.946	0.018	0.003	0.912	0.982
Wanted last child										
Yes*										
No	0.961	0.016	0.016	0.930	0.993	0.961	0.016	0.016	0.930	0.992
Marital status										
Married*										
Not Married	1.017	0.020	0.402	0.978	1.057	1.014	0.020	0.478	0.975	1.055
Employment status										
Unemployed*										
Employed	1.024	0.013	0.059	0.999	1.050	1.024	0.013	0.059	0.999	1.050
Husband educ. level										
No education*										
Primary	1.041	0.020	0.031	1.004	1.080	1.042	0.019	0.028	1.005	1.081
Secondary	1.080	0.020	0.000	1.041	1.120	1.083	0.020	0.000	1.044	1.122
Higher	1.093	0.024	0.000	1.047	1.140	1.096	0.024	0.000	1.051	1.143
Husband employment										
Unemployed*										
Employed	0.933	0.044	0.138	0.851	1.023	0.937	0.044	0.165	0.855	1.027
Inflation zero group	Coef.	SE	P-value	95%	CI	Coef	SE	P-value	95%	CI
Constant	0.833	0.250	0.001	0.342	1.324	0.780	0.262	0.003	0.265	1.294
Age group										
15 – 24*										
25 - 39	-0.115	0.057	0.042	-0.226	-0.004	-0.105	0.060	0.079	-0.222	0.012
40 - 49	-0.179	0.082	0.030	-0.340	-0.017	-0.168	0.087	0.053	-0.337	0.002
Region										
North Central*										
North East	-0.330	0.071	0.000	-0.469	-0.191	-0.403	0.076	0.000	-0.552	-0.255
North West	0.652	0.068	0.000	0.519	0.785	0.622	0.072	0.000	0.481	0.763
South East	-1.045	0.143	0.000	-1.326	-0.763	-1.043	0.159	0.000	-1.353	-0.732
South South	1.293	0.084	0.000	1.128	1.459	1.396	0.090	0.000	1.221	1.571
South West	-0.463	0.104	0.000	-0.668	-0.258	-0.369	0.108	0.001	-0.581	-0.157
Residence										
Urban*										
Rural	0.504	0.059	0.000	0.388	0.619	0.550	0.064	0.000	0.426	0.675
Educational level										
No education*										
Primary	-0.684	0.059	0.000	-0.799	-0.568	-0.720	0.063	0.000	-0.843	-0.597
Secondary	-1.148	0.077	0.000	-1.299	-0.996	-1.206	0.084	0.000	-1.369	-1.042
Higher	-2.165	0.252	0.000	-2.658	-1.671	-2.565	0.358	0.000	-3.267	-1.864
Religion										
Christian*										
Islam	-0.168	0.069	0.016	-0.304	-0.032	-0.173	0.074	0.019	-0.318	-0.029
Traditional/Other	0.692	0.178	0.000	0.344	1.041	0.706	0.184	0.000	0.345	1.067
Wealth index										
Poorest*										
Poorer	-0.585	0.050	0.000	-0.683	-0.487	-0.607	0.052	0.000	-0.709	-0.505
Middle	-1.164	0.062	0.000	-1.285	-1.042	-1.198	0.065	0.000	-1.326	-1.069
Richer	-1.579	0.082	0.000	-1.738	-1.419	-1.608	0.086	0.000	-1.777	-1.439

Richest	-2.111	0.128	0.000	-2.361	-1.861	-2.191	0.142	0.000	-2.469	-1.914
Parity										
1*										
2 - 4	0.189	0.065	0.004	0.062	0.317	0.182	0.069	0.008	0.048	0.317
More than 4	0.252	0.077	0.001	0.100	0.403	0.249	0.081	0.002	0.090	0.409
Wanted last child										
Yes*										
No	-0.109	0.077	0.161	-0.261	0.043	-0.129	0.083	0.119	-0.290	0.033
Marital status										
Married*										
Not Married	0.115	0.085	0.177	-0.052	0.283	0.108	0.090	0.227	-0.067	0.284
Employment status										
Unemployed*										
Employed	-0.357	0.043	0.000	-0.441	-0.274	-0.372	0.045	0.000	-0.460	-0.284
Husband educ. level										
No education*										
Primary	-0.687	0.056	0.000	-0.797	-0.576	-0.714	0.059	0.000	-0.830	-0.597
Secondary	-0.652	0.061	0.000	-0.772	-0.533	-0.661	0.064	0.000	-0.787	-0.535
Higher	-0.913	0.095	0.000	-1.099	-0.728	-0.945	0.103	0.000	-1.147	-0.742
Husband employment										
Unemployed*										
Employed	-0.228	0.223	0.307	-0.665	0.209	-0.224	0.233	0.337	-0.680	0.233
alpha	0.173	0.004		0.166	0.182	0.174	0.004		0.166	0.182
ln(alpha)	-1.752	0.024	0.000	-1.798	-1.705	-1.751	0.024	0.000	-1.798	-1.705

*reference category   SE = standard error   CI = confidence interval

4. Discussion

The present study compared the performance of the ZIP, ZINB, HP and HNB models in handling excess zeros associated with the number of ANC visits, a measure of ANC utilisation among women of reproductive age in Nigeria. These regression models are the most prominent and effective models not only for handling excess zeros in a count data but also for accommodating over dispersion resulting from the variance being greater than the mean (Staub & Winkelman, 2013). Although previous studies in Nigeria (Awusi et al., 2009; Dairo & Owoyokun, 2010; Onasoga et al., 2012) have examined the factors affecting ANC utilisation using the traditional regression methods which treat count data as a categorical variable (presence/absence), it is essential to address excess zeros emanating from the number of ANC visits using the model(s) that best describe it. Such exercise will not only provide the best model for identifying the factors associated with ANC utilisation, but will also set the pace for appropriate use of the model that addresses excess zeros in count dataset. To the best of our knowledge, this study represents the first effort to examine the underlying factors affecting ANC utilisation measured in terms of the number of ANC visits by women of reproductive age in Nigeria, using zero-inflated (ZIP and ZINB) and hurdle (HP and HNB) models.

In the present study, the socio-demographic characteristics of the participants showed that majority of the women were between 25-39 years, from the North West geo-political region, lived in the rural areas and practised Islam. Not surprising, women having these characteristics respectively had the highest proportion of zero ANC visits except for the age group 25-39 years with the least. Surprisingly, almost half of the participants had no education but majority of their husbands/partners were employed. Meanwhile, on the average, the study revealed that the women had the expected minimum number of ANC visits, as recommended by WHO (WHO, 2013; UNICEF, 2016). This supports a previous finding (Onasoga et al., 2012), from a study conducted at Ife Central, South West Nigeria, that majority of the women attended ANC regularly. As expected, the number of ANC visits evidently did not only show excess zeros but also showed over-dispersion in the dataset. The implication of this is that the zero-inflated (ZIP and ZINB) or hurdle (HP and HNB) models would be appropriate; however, the ZINB/HNB model was more appropriate for its robustness to over-dispersion compared with ZIP/HP model. This is in agreement with a finding (Ridout, et al., 2001) where ZINB was preferred to ZIP model on the basis that over-dispersion was still evident in an apple shoot propagation data.

Besides, the fact that the finding revealed a significant estimated dispersion parameter of value greater than zero further buttresses the appropriateness of ZINB/HNB model over the ZIP/HP model. Meanwhile, the estimated standard errors of the ZIP/HP model (not presented) being smaller than those of the ZINB/HNB model also explains why ZINB/HNB model led to some regression parameters being statistically insignificant. This shows that, in the presence of over-dispersion, the ZIP/HP regression model underestimates the standard error and biases the parameter estimates by overstating the

significance of the regression parameters. This result agrees with the study carried out by previous researchers (Fang, 2008; Ismail & Zamani, 2013; Al Mamum, 2014; Muniswamy et al., 2015). These existing studies reported that in the presence of over-dispersion and excess zeros, the ZINB model fits better than a ZIP model. Besides, the summary and fit statistic (AIC /BIC) attested to the fact that ZINB was more appropriate than other (ZIP, HP or HNB) regression models in handling excess zeros observed in the number of ANC visits. In addition, in agreement with the findings of Muniswamy et al. (2015) in which Likelihood ratio test, Wald test and Score test were used to compare the ZIP and ZINB models, ZINB was found to be a better model. Meanwhile, the results also suggested that HNB could be used to model the number of ANC visits. As opined by He et al. (2014), this could be possible if the sources (women who never visit and those who could but did not visit) of observed excess zeros were clearly distinguished from one another.

Findings from this study showed that age was a significant predictor of ANC utilisation rate in which older women visited ANC facilities more than younger ones. This finding agrees with a study carried out in south west Nigeria (Dairo & Owoyokun, 2010) in which older women were found to be more likely to utilise ANC compared to the younger women. On the contrary, the result disagrees with a study conducted in south south Nigeria (Awusi et al., 2009) which reported that older women utilised ANC less than younger women. Geopolitical zones were also found to be associated with ANC utilisation though with women in the southern region utilising ANC more compared to women in the northern region. A previous study, Mugo et al. (2010) has also documented this similar finding. While place of residence determined the likelihood of not untilising ANC, it was not associated with the rate of ANC utilisation. This finding, however, disagrees with previous studies (Dairo & Owoyokun, 2010; Adamu, 2011; Onasoga et al., 2012) where urban women attended ANC more frequently than rural women. This may be because urban women were more informed than rural women about the importance of ANC utilisation, as suggested by some previous studies conducted in Nigeria (Dairo & Owoyokun, 2010; Adamu, 2011). Parity was found to have a significant negative effect on ANC utilisation; respondents with higher parity tend to utilise ANC less than women with lower parity. This finding is similar to previous reports from a research conducted in south south Nigeria (Awusi et al., 2009) and a systematic review carried out in developing countries (Simkhada et al., 2008). A possible explanation could be that those with higher parity had more experience and may not appreciate the importance of utilising ANC, according to Awusi et al. (2009). Increased level of maternal education and husband's/partner's education also had a significant positive effect on ANC utilisation. In agreement with this submission, previous studies have reported similar findings (Nisar & White, 2003; Simkhada et al., 2008; Awusi et al., 2009).

In line with a similar study by Dairo and Owoyokun (2010), we found wealth quintile and religion to be associated with ANC utilisation such that women in the richest category and Muslims were the highest users. This might be because some Christian organizations have spiritual houses that offer care for pregnant women and so some of these women would rather visit such centres for ANC than come to the hospitals (Dairo & Owoyokun, 2010). Employed women had a significantly higher likelihood of utilising ANC. This observation is similar to that of a previous study on the determinants of urban-rural differential in ANC utilisation amongst women of reproductive age in Nigeria, using 2008 NDHS dataset (Babalola et al., 2014). According to the study, this might be because employed women may have more resources to access ANC services than those unemployed.

5. Conclusion

The Zero inflated negative binomial model was found to fit the data on number of antenatal care visits more than the Zero inflated Poisson, Hurdle Poisson and Hurdle negative binomial regression models in the presence of excess zeros. Age and "wanting last child" significantly influenced the rate of ANC utilisation, while rural-residence and being employed significantly determined the likelihood of not utilising ANC. In conclusion, in the analysis of count data with indistinguishable sources of excess zeros such as the number of ANC visits, the Zero inflated negative binomial regression model should be used.

References

Adamu, H. S. (2011). *Utilisation of maternal health care services in Nigeria: An analysis of regional differences in the patterns and determinants of maternal health care use* (Unpublished master's thesis). The University of Liverpool.

Al Mamun, M. A. (2014). *Zero-inflated regression models for count data: an application to under-5 deaths*(Unpublished master's thesis). Ball State University Muncie, Indiana.

Awusi, V. O., Anyanwu, E. B., & Okeleke, V. (2009). Determinants of antenatal care services in Emevor village, Nigeria. *Benin Journal of Postgraduate Medicine*, *11*(1), 21-26

Babalola, B. (2014). Determinants of urban-rural differentials in antenatal care utilisationin Nigeria. *African Population Studies*, *28*(3), 1263-1273

Cook, R., Dickens, B. M., Wilson, A. F, & Scarrow, S. E. (2001). *Advancing safe motherhood through human rights.* Geneva: World Health Organization. Retrieved from https://apps.who.int/iris/bitstream/10665/66810/1/WHO_RHR_01.5.pdf

Dairo, M. D., & Owoyokun, K. E. (2010). Factors affecting the utilisationof antenatal care services in Ibadan, Nigeria. *Benin Journal of postgraduate Medicine, 12*(1), 3-13. http://doi.org/10.4314/bjpm.v12i1.63387

Dalrymple, M. L., Ford, R. P., & Hudson, I. L. (2003). Finite Mixture, Zero-inflated Poisson and Hurdle models with application to SIDS. *Computational Statistics & Data Analysis, 41*, 491-504. http://doi.org/10.1016/S0167-9473(02)00187-1

Fang, R. (2008). *Zero-inflated negative binomial (ZINB) regression model for over-dispersed count data with excess zeros and repeated measures: an application to human microbiota sequence data* (Master's thesis, North China Coal Mining Medical College). Retrieved from https://dspace.library.colostate.edu/.../FANG_ucdenveramc_1639M_10037.pdf?seque

Greene, W. H. (1994). Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models. NYU Working Paper No. EC-94-10. Retrieved from https://ssrn.com/abstract=1293115

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

He, H., Tang, W., Wang, W., & Crits-Christoph, P. (2014). Structural zeroes and zero-inflated models. *Shanghai Archives of Psychiatry*, *26*(4), 236–242. http://doi.org/10.3969/j.issn.1002-0829.2014.04.008

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

Hu M. C., Pavlicova, M., & Nunes, E. V. (2011). Zero-inflated and Hurdle Models of Count Data with Extra Zeros: Examples from an HIV-Risk Reduction Intervention Trial. *Am J Drug Alcohol Abuse*, *37*(5), 367–375. https://doi.org/10.3109/00952990.2011.597280.

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. https://doi.org/10.1023/A:1021594923546

Ismail, N., & Zamani, H. (2013). Estimation of claim count data using negative binomial, generalized Poisson, zero-inflated negative binomial and zero-inflated generalized Poisson regression models. *Casualty Actuarial Society E-Forum*. Retrieved from https://www.casact.org/pubs/forum/13spforum/Ismail%20Zamani.pdf.

Kawungezi, P. C., AkiiBua, D., Aleni, C., Chitayi, M., Niwaha, A., Kazibwe, A., … Nakubulwa, S. (2015). Attendance and utilisationof antenatal care (ANC) services: Multi-center study in upcountry areas of Uganda. *Open Journal of Preventive Medicine, 5*(3), 132–142. https://doi.org/10.4236/ojpm.2015.53016

Lambert, D. (1992). Zero-Inflated Poisson Regression with an Application to Defects in Manufacturing. *Technometrics, 34*(1), 1-14. https://doi.org/10.2307/1269547

Lincetto, O., Mothebesoane-anoh, S., Gomez, P., & Munjanja, S. (2010). Antenatal Care: Opportunities for Africa's Newborns. New York: World Health Organisation. Retrieved from http://www.who.int/pmnch/media/publications/aonsectionIII_2.pdf

Lord, D. (2006). Modelling motor vehicle crashes using Poisson-gamma models: examining the effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter. *Accident Analysis & Prevention*, *38* (4), 751–766

Mugo, N. S., Dibley, M. J., & Agho, K. E. (2010). Prevalence and risk factors for non-use of antenatal care visits:

analysis of the 2010 South Sudan household survey. *BMC Pregnancy and Childbirth, 15,* 68. https://doi.org/10.1186/s12884-015-0491-6

Mullahy, J. (1986). Specification and testing of some modified count data models. *J Econometrics, 33*, 341–365. https://doi.org/ 10.1016/0304-4076(86)90002-3

Muniswamy, B., Molla, D. T., & Reddy, N. K. (2015). Comparison of test statistic for zero-inflated negative binomial against zero-inflated Poisson model. *Indian Journal of Science and Technology*, *8*(4), 349–357. https://doi.org/10.17485/ijst/2015/v8i1/59610

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

Nisar, N., & White, F. (2003). Factors affecting utilisationof Antenatal Care among reproductive age group Women (15-49 years) in an urban squatter settlement of Karachi. *Journal of Pakistan Medical Association, 53*(2), 47-53.

Onasoga, O. A., Afolayan, J. A., & Oladimeji, B. D. (2012). Factors influencing utilisationof antenatal care services among pregnant women in Ife Central LGA, Osun State, Nigeria. *Advances in Science and Research, 3,*1309-1315

Ridout, M., Hinde, J., & Demetrio, C. (2001). A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. *Biometrics, 57*(1), 219-223.

Sedgh, G., Singh, S., & Hussain, R. (2014). Intended and unintended pregnancies worldwide in 2012 and recent trends. *Studies in Family Planning*, *45*(3)*,* 301–314. https://doi.org/ 10.1111/j.1728-4465.2014.00393.x.

Simkhada, B., Teijlingen, E. R., Porter, M., & Simkhada, P. (2008). Factors Affecting the Utilisationof Antenatal Care in Developing Countries: Systematic Review of the Literature. *Journal of Advanced Nursing*, *61*, 244–260. https://doi.org/ 10.1111/j.1365-2648.2007.04532.x.

Stata Statistical Software (2011). StataCorp LP (12). 4905 Lakeway Drive College Station, Texas 77845 USA.

Staub, k., & Winkelmann, R. (2013). Consistent estimation of zero-inflated count models. *Health Economics, 22*, 673–686. https://doi.org/10.1002/hec.2844

United Nations International Children's Emergency Fund (2016). *UNICEF Data: monitoring the situation of children and women*. Retrieved from http://data.unicef.org/topic/maternal-health/antenatal-care

Vuong Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica, 57,* 307–333. https://doi.org/ 10.2307/1912557

World Health Organisation (2015). Trends in maternal mortality: 1990 to 2015: estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division, Geneva. Retrieved from http://www.afro.who.int/sites/default/files/2017-05/trends-in-maternal-mortality-1990-to-2015.pdf.

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

Xu, L., Paterson, A. D., Turpin, W., & Xu, W. (2015). Assessment and Selection of Competing Models for Zero-Inflated Microbiome Data. *PLoS ONE,* 10(7), e0129606. https://doi.org/10.1371/journal.pone.0129606

Yusuf, O. B., & Ughali, L. (2015). 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. https://doi.org/ 10.5923/j.ajms.20150503.04

Copyrights