Application of Count Regression Models for Factors of Antenatal Care (ANC) Utilization in Ethiopia

Zemene Yohannes[1] Birhanu Woldeyohannes[2]
1.Assistant Researcher in Industrial Development Policy Studies Centre, Ethiopian Policy Studies Institute (PSI)
2.Technical and Financial Team Leader, Addis Ababa Transport Bureau, Ethiopia
*Email: zemeneyohannes@gmail.com/ abukamsc2@gmail.com

Abstract
Adequate antenatal care is necessary to improve the health status of peoples in a country as it contributes the reduction of infant and maternal mortalities. The coverage and the quality of antenatal care plays significant role in maternal health. Ethiopia has witnessed an encouraging improvement in the ANC coverage (at least one visit) in the last decade from 62% in 2016 to 74% in 2019. But only 43% of women had at least four ANC visits during their last pregnancy. The purpose of this study is to identify factors that affect antenatal care visits based on 2019 Ethiopia Mini Demographic and Health Survey (EMDHS) dataset using count regression model. A total of 8,885 women aged 15-49 years were asked for the EMDHS-2019 out of which 3,927 women were illegible and asked for their ANC experiences for their last pregnancy. This empirical study aimed at investigating the impacts of the socioeconomic and demographic characteristics on the utilization of antenatal care. To investigate the impacts of the factors on the number of ANC visits, count models such as Poisson, Negative Binomial (NB), Zero Inflated Poisson (ZIP) and Zero-inflated negative binomial (ZINB) models were examined to select the best fitted model. The selection criterion Akaike information criterion (AIC), Bayesian information criterion (BIC) and Log Likelihood were applied to select the best model. The ZIP model was found to be the best fitted model for the data. Based on ZIP model, the explanatory variables region, age, place of residence, education, religion, wealth index and marital status were statistically significant factors on the frequency of ANC visits. Hence, efforts are still required to put on creating and expanding awareness on the utilization of ANC. The policy makers along with both private sectors and civil societies have to facilitate the increase of awareness and the experiences of the utilization of ANC in consideration of households within different groups of a society.
Keywords: EMDHS-2019, antenatal care, Count Regression model
DOI: 10.7176/JHMN/101-02
Publication date: June 30th 2022

1. Introduction

Maternal mortality is sensitive indicator of inequality. Information about the levels and trends of maternal mortality is needed not only for what it tells us about the risks of pregnancy and childbirth, but also for what it implies about women's health in general, and their social and economic status. Thus, maternal mortality is not just a "health disadvantage"- it is also a "social disadvantage" (WHO, 1986; BS Garg et al, 2006). For Maternal mortality, the World Health Organization (WHO) defines a maternal death as the death of a woman during pregnancy or within 42 days after its end regardless of duration or place of the pregnancy. Maternal health is a useful indicator to assess not only women's health status but also the accessibility, sufficiency, and effectiveness of country's health service system.

Over the last two decades, the world made substantial progress in reducing maternal mortality. The global Maternal Mortality ratio (MMR) in 2017 is estimated at 211 maternal deaths per 100,000 live births, representing a 38% reduction since 2000, when it was estimated at 342. The Ending Preventable Maternal Mortality (EPMM) target for reducing the global MMR by 2030 was adopted as SDG target 3.1: reduce global Maternal Mortality Rate (MMR) to less than 70 per 100,000 live births by 2030 [14]. Sub-Saharan Africa remains the region with the highest maternal mortality rate in the world. In 2017, the region had Sub-Saharan Africa and Southern Asia accounted for approximately 86% (254,000) of the estimated global maternal deaths in 2017 with sub-Saharan Africa alone accounting for roughly 66% (196,000) [13].

The 2019 EMDHS results show that 74% of women age 15-49 with a live birth in the 5 years before the survey received antenatal care (ANC) from a skilled provider for their most recent birth. 43% of women had at least four ANC visits during their last pregnancy. The proportion of women age 15-49 who received ANC from a skilled provider has increased over time, from 28% in 2005 and 34% in 2011 to 62% in 2016 and 74% in 2019.

ANC from a skilled provider varies by mother's age, from a high of 77% among women age 20- 34 to a low of 59% among women age 35-49. ANC from a skilled provider is lowest for sixth or higher-order births (58%) and highest for first-order births (83%). Urban women are more likely than rural women to receive ANC from a skilled provider (85% and 70%, respectively). Regionally, ANC coverage from a skilled provider is highest in Addis Ababa (97%) and lowest in Somali (30%). ANC from a skilled provider increases with increasing mother's education, from 62% among women with no education to nearly 100% among women with more than a secondary education. Women in the highest wealth quintile (95%) are more likely than those in the lowest quintile (47%) to receive ANC from a skilled provider (EMDHS 2019).

EMDHS 2019 revealed that the there is an increment in the frequency of ANC visits and a decline of maternal mortality. For instance, the proportion of women age 15-49 who received ANC from a skilled provider has increased over time, from 28% in 2005, 34% in 2011, 62% in 2016 and 74% in 2019. Even if the frequency of ANC visits is increased from time to time, it remains the lowest to be among in the world. The focus of this study was investigating the predicting factor of the frequency of ANC visits in Ethiopia. Therefore, the results of this study will provide prominent information on the type and size of significantly affecting factors of ANC in order to fill the gaps on preventing maternal deaths' existed in the country.

A study on antenatal care utilization and associated factors in West Shewa, Oromiya region, Ethiopia shows that age, educational status of respondents, parity, household distance from health post, and participation in decision making were statistically associated with ANC service utilization from rural Health Extension workers [3].

The type of residence, average annual income, level of the mothers, mothers having some education, and walking distance from a health facility were also significantly associated with being able to receive ANC during pregnancy in the study of determinants of equity in utilization of maternal health services in Butajira, Southern Ethiopia [7].

A study was conducted on antenatal care service utilization and associated factors in Metekel Zone, Northwest Ethiopia. Place of residence, educational status, husband's educational status, possessing radio, monthly income and knowledge about antenatal care were found to have a statistically significant association with antenatal care service utilization [5].

The main purpose of this study is to identify factors that affect the number of ANC visits. The study uses EMDHS-2019 data, on which 3,927 women from all regions. This study applied the count models: Poisson, NB, ZIP and ZINB count regression models. The criterion AIC, BIC and Log Likelihood were applied to select the best fit model for the data used.

2. Methodology

This study aimed at investigating the impact and the major socioeconomic and demographic characteristics on the incidences of ANC visits by applying count regression models using STATA and SPSS tools for data analysis. The EMDHS-2019 was applied to investigate the covariates of ANC visits in Ethiopia.

2.1. Count regression models

Count event is the number of times that an event occurred in a specified place and time. The count dependent variable has nonnegative integer response values $(0, 1, 2 \ldots)$ to which to be explained in terms of set of covariates (X). These models can be applied to examine the occurrence and frequency of events.

2.2. Poisson regression model (PRM)

Poisson regression model is the starting point to analyse count data. It can be used to model the number of occurrences of an event of interest. The Poisson probability density function has a single parameter$\boldsymbol{\lambda} > \mathbf{0}$. This parameter is mean incidence rate of a rare event$\boldsymbol{y}$.

Generally, we can investigate a Poisson dependent variable to see the impact of $X_i\beta$ on the probability of the occurrence response variable. The vectors $\boldsymbol{Xi}$; $\boldsymbol{i} = \mathbf{1}, \mathbf{2}, \ldots, \boldsymbol{p}$ are the explanatory variables and the $\beta_{(p+1)}Xi$ is a vector of $(\mathrm{p} + 1)$ parameters where the intercept $\boldsymbol{\beta_0}$ and the coefficients for the $\boldsymbol{p}$ covariates $\beta 1, \beta 2, \ldots, \beta p$. The objective is to estimate the coefficients; $\boldsymbol{\beta's}$ to interpret the result. Thus we declare the count dependent variable $\boldsymbol{Y_i}$ as

$$Y_i = Poisson[\exp(X_i\beta)] \qquad (1)$$

The Poisson regression assumes that the observed counts are generated from a Poisson distribution which is given by:

$$p(Y = y/\lambda) = \frac{e^{-\lambda}\lambda^y}{y!} \quad ; y = 0,1,2, \ldots \qquad (2)$$

where, $\boldsymbol{Y}$ is a random variable and the values y's are the possible outcomes or the realizations of the random variable $\boldsymbol{Y}$. Within a specified time interval, the Poisson response variables defined in $N = \{0,1,2, \ldots\}$ represents number of events occurred. The Poisson distribution is characterized with the property that the mean and the variance are assumed to be equal to λ. That is

$$E(Y_{poisson}) = Varpoission(Y) = \lambda \qquad (3)$$

The possible value of $\boldsymbol{\lambda}$ is clearly a positive number but not necessarily an integer. This characteristic of the distribution is described as "equi-dispersed". The mean of a random variable, $E(Y_i)$, with Poisson distribution shown in Equation (1) above can be given as:

$$E(Y_i) = E\{Piosson[\exp(X_i\beta)]\} \quad (4)$$

From the property of the Poisson distribution it follows that

$$\lambda_i = \exp(X_i\lambda) = \exp(\beta_o + \beta_1 x_1 + ... + \beta_p x_p) \quad (5)$$

Due to the canonical log-link functions property described in Equation (6), the Poisson regression models are also called Poisson log-link or simply log-linear models. The Poisson distribution is one of the exponential classes of probability distribution functions. Thus, the generalized linear model approach can be used to relate the Poisson response to the explanatory variables. Given p explanatory variables and taking the natural logarithms of the Equation (5) we have the following:

$$\ln \lambda_i = \exp(\beta_o + \beta_1 x_1 + ... + \beta_p x_p) \quad (6)$$

From this we can say that the link function for Poisson model is the natural logarithm. In addition, we can understand that the Poisson model is a log-linear model. The Poisson regression model is the benchmark model for count data in much the same way as the normal linear model is the benchmark for real-valued continuous data. The Poisson model is simple, and it is robust [12].

The most popular model for count data is the Poisson model, which is based on the property that the mean and variance of the dependent variable is assumed to be equal [6]. However, this is not always the case, as the variance sometimes exceeds the mean. Over dispersion can be modelled by using negative binomial (NB) regression model, but more models accounting for over dispersion exist.

2.3. Negative Binomial Model (NBM)

Under the PRM, the observed difference (heterogeneity) among the sample members is specified by the rate λ_i as a function of observed $Xi's$. In practice, the PRM underestimates the amount of dispersions in the observed data. Hence, the negative binomial regression model solves this failure by using the unobserved heterogeneity parameter α.

By adding an error term $\mathcal{E}$ that is assumed to be uncorrelated with the covariates $\boldsymbol{X's}$ in the relationships of Poisson distribution described in Equation (5), for p covariate we have the following resulting equations.

$$\mu_i = \exp(\beta_o + \beta_1 x_1 + ... + \beta_p x_p + \varepsilon_i) \quad (7)$$

Equivalently,

$$\mu_i = \exp(\beta_o + \beta_1 x_1 + ... + \beta_p x_p)\exp(\varepsilon_i) \quad (8)$$

Let $\delta_i \equiv \exp(\varepsilon_i)$, then

$$\mu_i = \exp(\beta_o + \beta_1 x_1 + ... + \beta_p x_p)\delta_i \quad (9)$$

In a linear regression model $E(\varepsilon_i) = 0$. This implies$E(\delta i) = 1$. From this assumption and applying Equation (7), $E(\mu\, i)$ is given as:

$$E(\mu_i) = E[xp(\beta_o + \beta_1 x_1 + ... + \beta_p x_p) + \varepsilon_i] \quad (10)$$

which is equivalent to

$$\begin{aligned} E(\mu_i) &= E[xp(\beta_o + \beta_1 x_1 + ... + \beta_p x_p)\delta_i] \\ &= E[\exp(\beta_o + \beta_1 x_1 + ... + \beta_p x_p]E(\delta_i) \end{aligned} \quad (11)$$

Finally, we have

$$\lambda_i X1 = \lambda_i$$

Therefore; under the assumptions of the NBRM, the expected rates for a given level of the independent variables will be the same in both PR and NB regression models. In other words the mean structure of the PRM and the NBRM is the same. However; the standard errors in the PRM will be biased downward, resulting in apparent large z-values and apparent small p-values [1].

Since the distribution of $\delta = exp(\varepsilon)$ is unknown, it is difficult to define the probability distribution of y/x. Thus this can be solved by introducing the gamma distribution for δ [9]. For count response variable conditional with the covariates X the negative binomial distribution is given by the following equation.

$$f(Y_i = y_i / x_i) = \frac{(\alpha\mu_i)^{y_i}\Gamma(\frac{1}{\alpha} + y_i)}{y_i\Gamma(\frac{1}{\alpha})(\alpha\mu_i + 1)^{\frac{1}{\alpha}+y_i}} \qquad (12)$$

The conditional NB distribution has conditional mean

$$E(Y = y_i / X) = \mu_i \qquad (13)$$

and the conditional variance is

$$\mathrm{var}(Y = y_i / X) = \mu_i(1 + (\frac{1}{\theta})\mu_i \qquad (14)$$

For $\alpha = \frac{1}{\theta}$, the conditional variance of the NB distribution is equivalent to $\mu_i(1 + \alpha\,\mu_i)$. If α converges to zero the NB distribution converges to Poisson distribution. In other words, the NB regression is the extension of Poisson that could collapse into Poisson regression with the dispersion parameter equal to zero. Due to this property, the Poisson and the NB are nested distributions. This important property provides the model comparison between Poisson and NB regressions. Therefore; negative binomial (NB) regression model is a generalization of the Poisson regression model. The NB regression model allows for over-dispersion by introducing an unobserved heterogeneity term for observations. Observations are assumed to differ randomly in a manner that is not fully accounted for by the observed covariates.

2.4. Over-dispersions

Equality of the mean and variance will be referred to as the equi-dispersion property of the Poisson. In real life situations this property is frequently violated. Thus, there is an over-dispersion or otherwise under-dispersion in the data. Over-dispersion is thus due to the variance exceed the mean of the response variable. Over-dispersion is closely related to the presence of unobserved inter-individual heterogeneity, but it can also arise from occurrence dependence between events. Zero event counts are often dominant, leading to a skewed distribution. Also, there may be a great deal of unobserved heterogeneity in the individual experiences of the event in question. Unobserved heterogeneity leads to over-dispersion; that is, the actual variance of the process exceeds the Poisson variance even after repressors are introduced [1].

2.5. Zero-Inflated Count Regression Models

The NB model might not be appropriate if the over-dispersion is caused by an excessive number of zeros in the outcome. In this case, alternative models such as zero inflated models are recommended. A ZIP model will reflect data accurately when over-dispersion is caused by excess of zeros. If over dispersion is attributed to factors beyond the inflation of Zeros, a ZINB model is appropriate [5].

2.6. Zero-Inflated Poisson (ZIP) Regression Model

The zero-inflated Poisson (ZIP) regression model is a modification of the familiar Poisson regression model that allows for an overabundance of zero counts in the data. The restrictions of the Poisson distribution which is the equality of the mean and the variance is violated because of a data are often over-dispersed. Thus the Poisson distribution underestimates the dispersion of the observed counts. The over-dispersion occurs when the single parameter λ of Poisson distribution is unable to fully describe event counts. Hence, this Poisson regression model in this regard is inadequate due to excess zeroes.

In some applications the proportion of the observed zeroes could exceed largely the proportion of the zeroes reproduced by the fitted Poisson distribution. In many cases the proportion of the zeroes will not affect the precision of the estimations. However; greater proportion of zeroes affects the estimations in Poisson regression models. Thus, Zero-Inflated Poisson regression model can be taken for handling the occurrences of a higher proportion of zeroes than the parent Poisson distribution.

Lambert (1992) introduced the zero inflated Poisson (ZIP) model in which $\lambda i = \lambda(xi;\beta)$ and the probability ωi is parameterized as a logistic function of the observable vector of covariates zi [8]. The occurrences of response are given by [1]

$$y_i = 0 \qquad \text{with probability } \omega_i$$

$$y_i = P[\lambda_i] \qquad \text{With probability } 1 - \omega_i \qquad (15)$$

The probability of the count random variable $Y_i = y_i$ is given by

$$P(Y_{yi}|x_i, z_i) = \begin{cases} \varphi(\gamma' z_i) + (1 - \varphi(\gamma' z_i))g(0|x_i); & where\ y_i = 0 \\ (1 - \varphi(\gamma' z_i))g(0|x_i); & where\ y_i > 0 \end{cases} \quad (16)$$

Lambert (1992) introduced the zero-inflated Poisson (ZIP) model in which $\lambda_i = \lambda(x_i; \beta)$ and the probability ω_i is parameterized as a logistics function of the observable vector of covariates z_i [12]. The occurrences of the responses are given by

$$\begin{cases} y_i = 0 & \text{with probality } \omega_i \\ y_i = P[\lambda_i] & \text{with probability } 1 - \omega_i \end{cases} \quad (17)$$

With the probabilities

$$\omega_i = \frac{e^{Z_i \gamma}}{1 + e^{Z_i \gamma}} \quad (18)$$

where $P\ [\lambda i]$ is a Poisson probability density function and if the probabilities ωi, zero-inflated Poisson distribution approaches to Poisson(λ). By subtracting from one both sides in the Equation (18) the following equation can be obtained.
With the probabilities

$$1 - \omega_i = \frac{1}{1 + e^{Z_i \gamma}} \quad (19)$$

the ZIP model assumes that some zeroes occur by a Poisson process and others were not even eligible to have the event occur. Some explanatory variable may have a significant effect to produce zero counts to the response variable. But it is not possible to identify which zero count is due to either of the processes. The essential idea is that the data come from two processes. In the first process, the outcome is always a zero count, while in the second the counts follow a standard Poisson process [4]. Thus, the two processes of the ZIP models for generating its count data can be described by the following probability distribution.

$$\begin{aligned} pr(y_i = r) &= \omega_i + (1 - \omega_i) e^{\lambda_i}; \qquad \text{where} \quad r = 0 \\ &= (1 - \omega_i) \frac{e^{\lambda_i} \lambda^r}{r!}; \qquad r = 1,2,3,\ldots \end{aligned} \quad (20)$$

The mean and the variance of the ZIP distribution are, respectively:

$$Ezip(Y_i) = (1 - \omega_i)\lambda_i \quad (21)$$

and

$$\operatorname{var}_{zip}(y_i) = (1 - \omega_i)(\lambda_i / \lambda_i^2)[(1 - \omega_i)\lambda_i^2] \quad (22)$$

The variance of the ZIP distribution the above Equation (22) can be equivalently expressed as:

$$\operatorname{var}_{zip}(y_i) = (1 - \omega_i)(\lambda(i) + \omega_i \lambda_i^2) \quad (23)$$

Hence the variance of the ZIPM is the quadratic function of its mean [10]. Because of the presence of over-dispersion due to the excess zeros in the data yields;

$$\operatorname{var}_{zip}(y_i) = (1 - \omega_i)(\lambda(i) + \omega_i \lambda_i^2) > \lambda_i = E_{zip}(y_i) \quad (24)$$

ZIP model fits simultaneously two separate regression models.

2.7. Zero-Inflated Negative Binomial (ZINB) Regression Model

For count data that are skewed with highly right tail, the NB distribution is appropriate instead of the Poisson distribution. Furthermore the zero inflated negative binomial (ZINB) regression appropriate to model over-dispersed data with an excess of zeros [11].
The probability distribution function of the ZINB regression model is given by:

$$pr(Y_i = y_i|x_i) = w_i + (1 - w_i)(1 + \alpha\lambda_i)^{-\frac{1}{\alpha}}$$

$$= (1 - w_i) \frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{y_i!\,\Gamma\left(\frac{1}{\alpha}\right)} \frac{(\alpha\lambda_i)^{y_i}}{(1 + \alpha\lambda_i)^{y_i + \frac{1}{\alpha}}}$$

Where i and are respectively the parameter of the negative binomial distribution and arbitrary over-dispersion parameter [12]. The mean and the variance of the ZINB are respectively given by

$$E(Y)_{ZINB} = (1-\omega_i)\lambda \equiv \lambda^*$$

Therefore; the variance of the ZINB distribution is a quadratic function of its mean. [10]
Similar to PR and ZIP models, as the dispersion parameter equal to zero the ZINBM becomes the NBM. Thus, NB and ZINB are nested models. This property is important for comparisons among the two nested groups.
AIC and BIC are used for model comparison. When one model is a special case of another, we can test the null hypothesis that the simpler model is adequate against the alternative hypothesis that the more complex model fits better [2].

3. Discussion of the Results

To understand how the different count regression models fit the antenatal care, the researcher examined count regression models for the frequency of antenatal care visits based on EMDHS, 2019 data. The Poisson, NB, ZIP and ZINB regression models were applied and examined to identify the determinants of the number of antenatal care visits. According to the model selection criteria, the ZIPRM was found an appropriate to this data (Table 1).

This study identified the socioeconomic and demographic determinants of the number of ANC visits based on EMIDHS, 2019 dataset. According to the results, mother's educational level has a positive relationship on ANC utilization. ANC visits increases with increase in mother's education level. Thus, with the opportunity of experiencing ANC for primary & secondary and above educational level were increased by 21.16% and 25.72%, respectively as compared to not educated mothers holding other variables in the model. According to the model selection criteria (AIC, BIC and Log Likelihood), the ZIP regression model is the most appropriate to fit this data (Table1). From the results, it was found that women who live in Addis Ababa region are found to have a higher opportunity to ANC and are 0.1107 times higher than Tigray (reference category). This could be as a result of higher health facilities and living standards in Addis Ababa, compared to other regions. The finding of the study on the variable religion has negative relationship on ANC visits revealed that the visits is significantly less than mother from orthodox. The study also revealed that wealth index is an important variable that affects the number of ANC visits. It has a positive effect. Amazingly, as wealth index increases the ANC visits also increases .From the results, it was found that mother who live in rural are found to have a decrease ANC vistits than Rural and are 0.1242 times lower than urban (reference category). In addition, mother who were in union more ANC visits than those who were never in union.

Table 1 Model Selection for Count Regression Model

Model	AIC	BIC	Log likelihood
Poisson	16715.09	16765.36	-8349.5431
NB	16405.9	16462.47	-8193.9523
ZIP	14658.49	14897.31	-7291.25
ZINB	16170.77	16239.89	-8074.383

Table 2. model results for number of ANC by ZIP model.

Covariate	Estimate	S.E	Z-value	P-value	IRR	95% CI of IRR	
						Lower	Upper
Intercept	.982269	.1536269	6.39	0.000	2.670509	1.97616	3.608778
Age	.0062999	.0017054	3.69	0.000	1.00632	1.002962	1.009689
Residence (Urban Ref.)							
Rural	-.1241766	.0339526	-3.66	0.000	.8832238	.8263619	.9439984
Region (Tigray, ref.)							
Afar	-.0766154	.0558915	-1.37	.070	.926246	.8301402	1.033478
Amahara	-.052677	.0400514	-1.32	.058	.9486864	.8770631	1.026159
Oromia	-.0625239	.0458123	-1.36	.072	.9393906	.8587184	1.027642
Somali	-.7420009	.0993059	-7.47	.000	.4761602	.3919438	.5784721
Benishangul	.0037569	.0443024	.08	.032	1.003764	.9202828	1.094818
SNNPR	-.1374725	.0492864	-2.79	.005	.8715583	.7913048	.9599511
Gambela	-.2446657	.0526055	-4.65	.000	.7829662	.7062609	.8680024
Harari	-.0774303	.0489985	-1.58	.114	.9254915	.840746	1.018779
Addis Ababa	.1107718	.0473882	2.34	.019	1.11714	1.018054	1.22587
Dire Dawa	-.0499727	.465797	-1.07	.082	.9512554	.8682572	1.042187
Education (no edu.)							
Primary	.1202429	.0254843	4.72	.000	1.127771	1.072824	1.185531
Secondary	.2116275	.0344588	6.14	.000	1.235687	1.154987	1.322026
Higher	.2572973	.039168	6.57	.000	1.29343	1.197851	1.396634
Religion(Orthodox)							
Catholic	-.243971	.153456	-1.59	.012	.7835103	.579993	1.058441
Protestant	-.0804936	.035915	-2.24	.025	.9227069	.8599889	.9899989
Muslim	-.063585	.0283803	-2.24	.025	.9383944	.887622	.992071
Traditional	-.2310529	.1818455	-1.27	.104	.7936975	.5557024	1.133621
Other	-2602269	.2283128	-1.14	.054	.7708767	.4927705	1.276693
Wealth(Poorest ref.)							
Poorer	.081999	.0370112	2.22	.027	1.085455	1.009503	1.16712
Middle	.0899022	.038154	2.36	.018	1.094067	1.015237	1.179019
Richer	.1277193	.038046	3.36	.001	1.136234	1.054588	1.224201
Richest	.1579252	.0440559	3.58	.000	1.171079	1.074201	1.276693
Marriage(Never in union)							
Married	.1767756	.13778	1.28	.099	1.193363	.9109492	1.563332
Living with partner	.2249289	.1709286	1.32	.088	1.252234	.8957579	1.750572
Widowed	.1839858	.1723029	1.07	.086	1.201999	.8755106	1.684878
Divorced	.1810814	.1461146	1.24	.015	1.198513	.9000562	1.595937
Separated	.1154012	.1556709	.74	.059	1.122324	.8272005	1.522739

4. Conclusions and recommendations

4.1. Conclusion

The purpose of this study was to identify, socioeconomic, demographic, health and environmental related determinants of the number of antenatal care visits in Ethiopia. In this study count regression model is used. From the exploratory results we could identify that there was an excess zeros. This implies that the standard Poisson model is not an adequate model to fit antenatal care visits. From the results of count regression models, zero inflated Poisson regression model is better fitted the data which is characterized by excess zeros and high variability in the non-zero outcome than any other models and therefore, it is selected as the best parsimonious model to predict the number of antenatal care visits in Ethiopia. For selected ZIP model, the explanatory variables such as age, residence, region, mother's education level, age, religion, wealth index and marital status are statistically significant factors on the number of antenatal care visits in Ethiopia.

4.2. Recommendations

Based on the findings of the study, we forward the following possible recommendations: The educational level of mothers plays an important role in frequency of ANC visits, efforts are needed to extend educational programmes aimed at educating mothers. Policies and programs aimed at addressing regional differences in ANC visits must be formulated and their implementation must be vigorously pursued. To achieve this, which have worked to some extent in the Somali and Gambela must be strengthened to achieve more results in the region, these measures could be extrapolated and applied in the remaining regions of the country. The concerned body should work closely with both the private sector and civil society to teach households to have sufficient knowledge and awareness on ANC.

References

1. Colin Cameron.A and Pravin K. Trivedi (1998). Regression analysis of count data. Econometric Society Monographs No. 30. Cambridge University Press, United Kingdom. Published in the United States by Cambridge University Press, New York.
2. Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Auto. Control* AC-19, 716723.
3. Birhanu Darega Gela, Nagasa Dida Bedada, Fikru Tafese Jaleta, Shimeles Ololo Sinkie (2014). Antenatal Care Utilization and Associated Factors from Rural Health Extension Workers in Abuna Gindeberet District, West Shewa, Oromiya Region, Ethiopia. *American Journal of Health Research*. Vol. 2, No. 4, pp. 113- 117. doi: 10.11648/j.ajhr.20140204.12.
4. David Giles (March, 2010). The Zero-Inflated Poisson Regression Model. Department of Economics, University of Victoria.
5. Gurmesa Tura (Jannuary-February 2007). Antenatal care service utilization and associated factors in Metekel zone. Benishangul Gumuz Region, North West Ethiopia.
6. Hoffmann JP. Generalized linear models: An applied approach. Pearson College Division, 2004.
7. Jemal Aliy, Damen Haile Mariam (2012). Determinants of equity in utilization of maternal health services in Butajira, Southern Ethiopia., Ethiop. *J. Health Dev*.2012; 26 Special Issue 1:265-270.
8. Lambert, D. (1992). Zero-In fated Poisson Regression Models with an Application to Defects in Manufacturing, *Techno metrics*, 34, 114.
9. Long, S.J. (1997). Regression Models for Categorical and Limited Dependent Variable. Sage Publications.
10. Minami, M., Lennert-Cody, C.E., Gao, W., Roman-Verdesoto, M. H. (2006). Modelling shark by catch: The zero-inated negative binomial regression model with smoothing. *ISM Research Memorandum 978.*
11. M. Minamia, C.E. Lennert-Cody ,W.Gao, M.Roman-Verdesoto (2007). Modelling shark by catch: The zero-inflated negative binomial regression model with smoothing. *Fisheries Research* 84 (2007) 210221.
12. Rainer Winkelmann (2008). Econometric Analysis of Count Data., Fifth edition. Springer-Verlag Berlin Heidelberg.
13. WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division (2019). Trends in Maternal mortality, 2019: Estimates developed by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division.
14. World Health Organization. Regional Office for the Western Pacific. (2016). Sustainable development goals (SDGs).