ZERO TRUNCATED POISSON REGRESSION MODEL FOR REPRODUCTIVE PATTERNS ON COUNT DATA

B. Muniswamy[1], M. V. Lavanya[2*]

•

[1,2]Department of Statistics, Andhra University, Visakhapatnam, India.
Corresponding Author Email: *lavanya.mv9@gmail.com

Abstract

The number of children ever born is an important measure for understanding fertility patterns, which impact demographic structures and population growth. The problem relates to the modeling of count data that includes the truncation of zero values, specifically focusing on women who have experienced childbirth at least once. This study analyzes the factors that influence the number of children ever born (CEB) among women aged 15 to 50 in Andhra Pradesh, utilizing data from the National Family Health Survey (NFHS-5) conducted from 2019 to 2021. The study used Zero-Truncated Poisson (ZTP) and Zero-Truncated Generalized Poisson (ZTGP) models to identify major determinants, including religion, kind of cooking fuel used, place of delivery, wealth, age, and fertility choices. The ZTP regression model was found to be the best model and identifies significant determinants such as religion, wealth, age, and fertility preferences. The results show that rural residence, Muslim faith, and older age groups are associated with higher CEB, while wealthier women tend to have fewer children. The study shows the importance of implementing focused reproductive health activities, specifically in rural regions, to manage population growth and enhance the health outcomes of both mothers and children.

Keywords: Number of children ever born, Fertility Patterns, Zero-Truncated Poisson Model, NFHS, Reproductive Health, Under-dispersion

I. Introduction

The number of children ever born (CEB) quantifies the total count of live births among women aged 15 to 50 [19]. CEB, a summary of birth histories, is a quantitative measure of all women's live births during their lifetime. The CEB is a significant factor in shaping global population trends [5,23]. Population growth is not only influenced by it, but it also plays a crucial role in shaping the demographic age distribution. Fertility is a major component of Demography, which has three primary categories and refers to the natural ability for reproduction. Evaluating fertility trends [6] and prospective opportunities is vital to economic and social planning, workforce accessibility, and advancement [2]. Examining variations in reproduction rates among Indian states based on socioeconomic and demographic factors indicates significant variety [7]. The fertility rate in India has had a gradual decrease over the years, reaching 2.47 in 2012, 2.41 in 2013, 2.31 in 2014, 2.29 in 2015, 2.27 in 2016, 2.2 in 2017, 2.18 in 2018, 2.11 in 2019, 2.05 in 2020, and 2.03 in 2021 children per woman. The National Family Health Survey 2019-2021 (NFHS-5) offers comprehensive data on India's population, health, and nutrition [18]. This study aims to evaluate the prevalence of CEB (Children Ever Born) and its determining factors among women of reproductive age in Andhra

Pradesh.

Counting data with an excess of zeros is common in various fields, including engineering, biomedical research, public health, demography [1], economics, and social science. The basic Poisson regression model is the best strategy for analyzing a random variable Y expressing counts with equal sample mean and variance [8,21]. Count data displays significant variability when the sample variance is either smaller or bigger than the sample mean and is categorized as under-dispersion or over-dispersion [4,11]. Several models [27] have been suggested to address these variations, such as the negative binomial model [13], extended Poisson model [8], hurdle Poisson model [3,22], and truncated models [10]. The Generalized Poisson model was designed to analyze family fertility [15,27] and injury data [31]. However, it is frequently seen that count data exhibits a low frequency of zeros and is under-dispersed, indicating the absence of zero inflation in fertility [25]. This study examines the impact of several socioeconomic and demographic characteristics, such as site of residence, kind of cooking fuel used, place of delivery, wealth index, marital status [26], and caste, on the outcomes of women not experiencing infertility. To obtain the fertility rate of women aged 15–50, truncate the zero values in the count variable of the dataset [20,30]. The model-building procedure employs data from the NFHS-5 survey, specifically focusing on 8087 women from Andhra Pradesh who had given birth at least once. This paper examines the zero-truncated Poisson (ZTP) model [12], which accounts for both over and under-dispersion, as well as the zero-truncated generalized Poisson (ZTGP) model [28], which accounts for under-dispersion.

This study used secondary data as its basis. The data for this investigation was gathered from the fifth round of the NFHS, the most extensive sample survey representing the entire nation [17]. The NFHS is an Indian dataset derived from the seventh phase of the Demographic and Health Surveys (DHS) Program, carried out under the supervision of the National Institute for Population Research and Training of the Ministry of Health and Family Welfare. The NFHS Subject Reports are concise summaries of secondary data analysis from the 1992–93 National Family Health Survey (NFHS) conducted in India [16]. The National Family Health Survey (NFHS) collected information from around 90,000 women in India, covering all aspects of demographics and health. The Indian Ministry of Health and Family Welfare conducted this survey, which provides in-depth information on maternal and child health, family planning practices, infant and child mortality, and the use of mothers' and children's services at the national and state levels [14]. IIPS performed the survey in collaboration with consultancy organizations and 18 population research centres across India. The East-West Center and Macro International, a U.S.-based consulting firm, offered technical help, while the United States Agency for International Development (USAID) provided financial support.

II. Methodology

2.1 Regression Models

The Poisson regression model is a widely used non-linear regression model for counting data [8]. Let Y represent the number of children ever born (CEB) to a woman of reproductive age in Andhra Pradesh. This variable follows a Poisson distribution, determined by the independent variables X_1, X_2, ..., X_{10}.

$$P(Y = y) = \frac{e^{-\mu}\mu^{y}}{y!} \quad ; for\ \mu > 0, y = 0,1,2, \ldots \tag{1}$$

It implies that μ is the exponential function of independent variables,

$$\mu = e^{\alpha+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_{10} X_{10}} = e^{x'\beta} \tag{2}$$

Here, α is the intercept, and β's are the Poisson regression coefficients.

For subject i,

$$ln(\mu_i) = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_{10} X_{10i} = x_i'\beta \quad (3)$$

where, $x' = [1 \quad X_1 \quad X_2 \ldots \quad X_{10}]$ (4)

Since the variable CEB represents a count and all observations are greater than zero, this study aimed to develop a regression model using the zero-truncated Poisson and zero-truncated generalized Poisson models [12].

2.1.1 Zero Truncated Poisson (ZTP) Model

For the Poisson distribution with the probability mass function (pmf) (1), the pmf for ZTP distribution is given by

$$P(Y = y|Y > 0) = \frac{\mu^y}{y![e^{\mu}-1]} \quad ; for\ \mu > 0, y = 1,2, \ldots. \quad (5)$$

The mean and variance of the ZTP random variable are as follows:

$$E(Y_i) = \frac{\mu e^{\mu_i}}{e^{\mu_i}-1}, \quad (6)$$

$$V(Y_i) = \frac{\mu e^{\mu_i}}{e^{\mu_i}-1}\left[1 - \left(\frac{\mu e^{\mu_i}}{e^{\mu_i}-1}\right)\right] \quad (7)$$

2.1.2 Zero Truncated Generalized Poisson (ZTGP) Model

Let a count response Yi ~ GP(μ,α); i =1, 2, …, n, then Yi has a probability function [29]:

$$f(y_i; \mu_i, \alpha|Y > 0) = \frac{1}{\left[exp\left(-\frac{\mu_i}{1+\alpha\mu_i}\right)-1\right]}\left[\frac{\mu_i}{1+\alpha\mu_i}\right]^{y_i}\frac{(1+\alpha y_i)^{y_i-1}}{y_i!}exp\left(-\frac{\alpha\mu_i y_i}{1+\alpha\mu_i}\right) \quad (8)$$

2.2 Accessing model adequacy and model comparisons

Subsequently, the loglikelihood, Akaike Information Criterion (AIC) [9], and Bayesian Information Criterion (BIC) were compared across all models to assess and choose the most appropriate model. The statistical tests were conducted using the professional statistical program R [32] and SPSS. The final model for the analysis was selected based on the greater loglikelihood and the minimum information criteria value [24].

III. Results

Table 1: *Descriptive statistics of the number of CEB*

Variable	N	Mean	Variance	Minimum	Maximum
CEB	8087	2.203	0.601	1	4

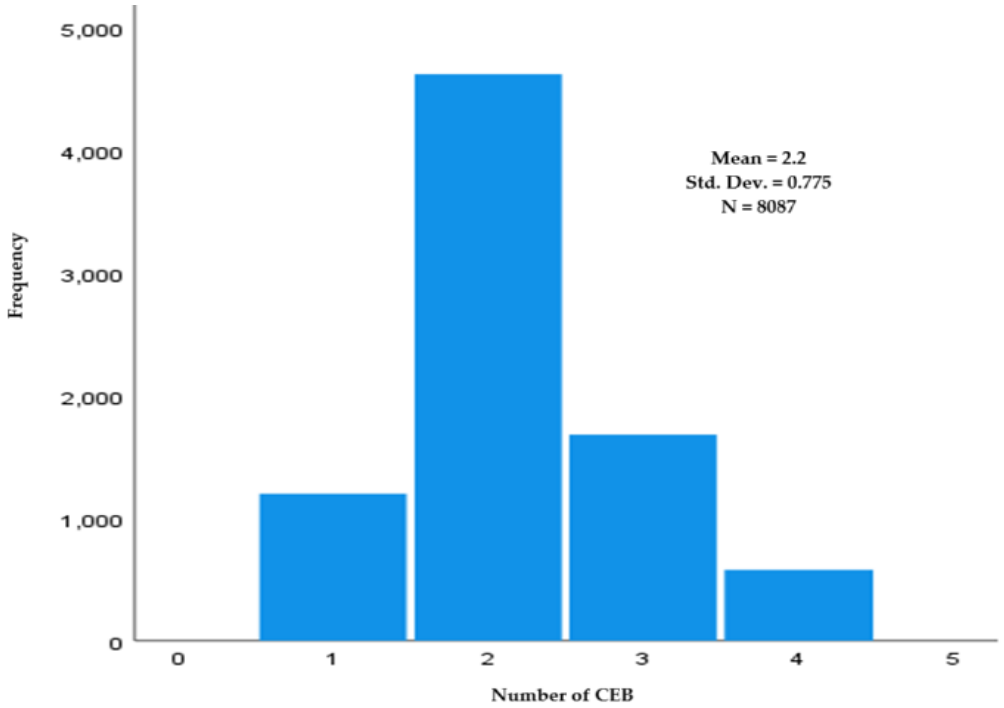

Figure 1: *Histogram of the Number of CEB*

Table 1 displays the descriptive statistics for the count of CEB, which is used as the response variable. The quantity of CEBs varies between 1 and 4. The dataset consisted of 8087 observations. The mean and variance of the number of CEB were calculated to be 2.203 and 0.601, respectively. These values indicate the data set exhibits under-dispersion. Figure 1 shows that the smallest number of children was 1, whereas the maximum number was 4.

Table 2: *Frequency distribution of CEB*

CEB	1	2	3	4
Frequency	1200	4627	1682	578
Percent	11.4	44.0	16.0	5.5

Table 2 displays the number of CEBs, their frequencies, and corresponding percentages. According to the table, 44% of women had two children, the most common number. In addition, 16% of women had three children, the second-highest percentage. Moreover, it is evident that in AP, the population of women with 2-3 children exceeded those with only one child and more than four children.

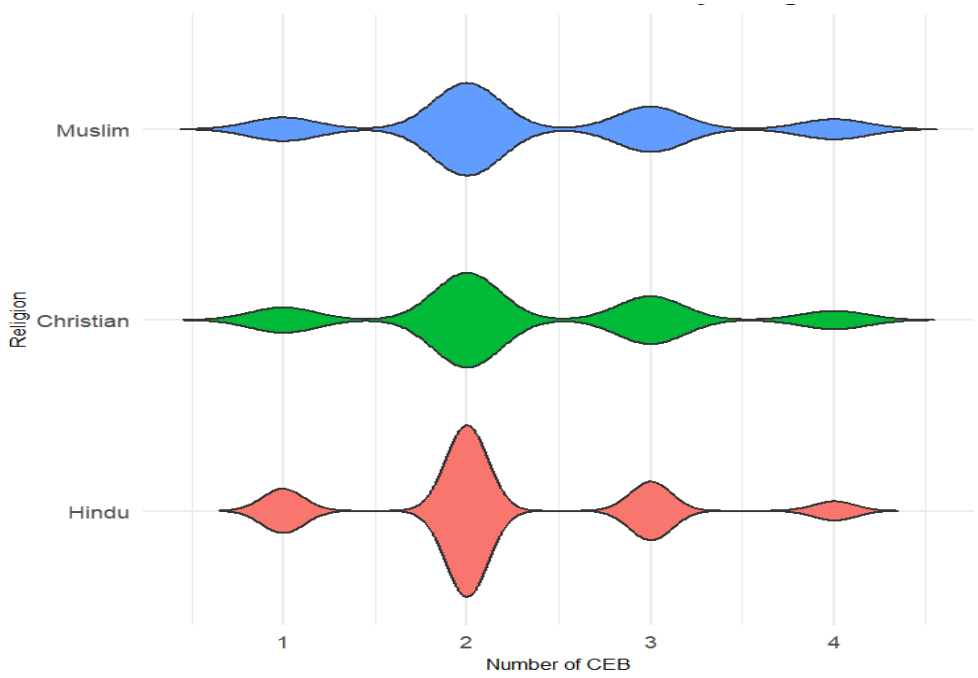

Figure 2: *Distribution of Number of CEB by Religion*

Figure 2 shows the distribution of the number of CEB among various religious groups. Most women had two children, corresponding to the maximum width of each violin plot, emphasizing the disparities in fertility rates among different religious groups.

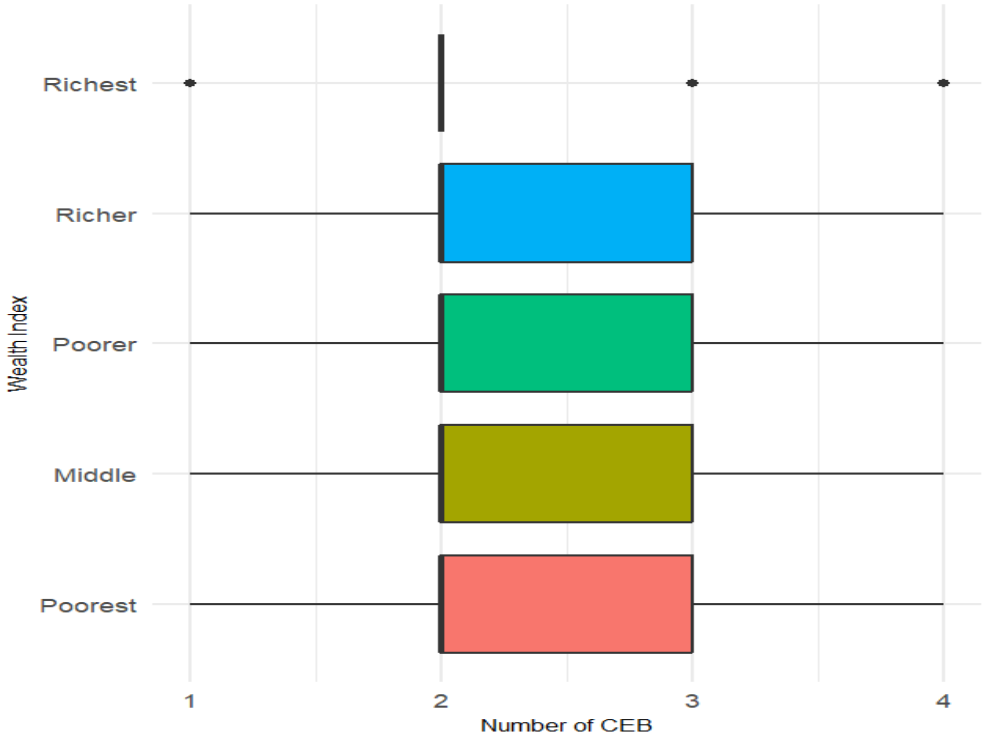

Figure 3: *Distribution of Number of CEB by Wealth Index*

Figure 3 demonstrates the relationship between the number of CEB and the wealth index. The median number of children decreases as we move from poorer to wealthier households.

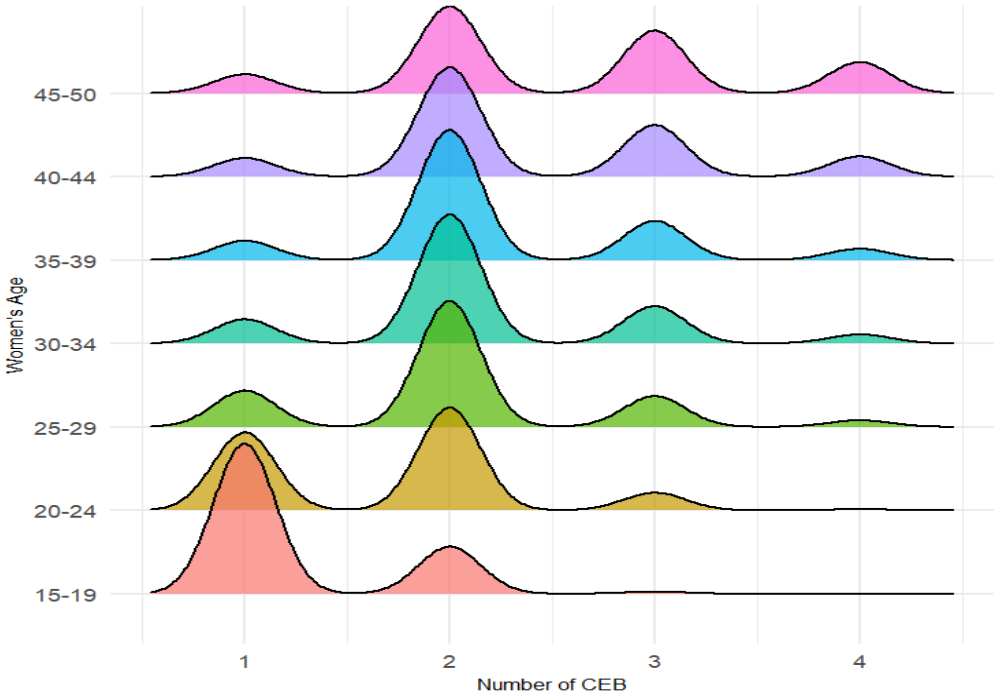

Figure 4: *Distribution of Number of CEB by Women's Age*

Figure 4 shows the distribution of the number of CEB across different age groups of women. It provides a clear visual representation of how fertility patterns differ by age, with older women generally having more children than younger women.

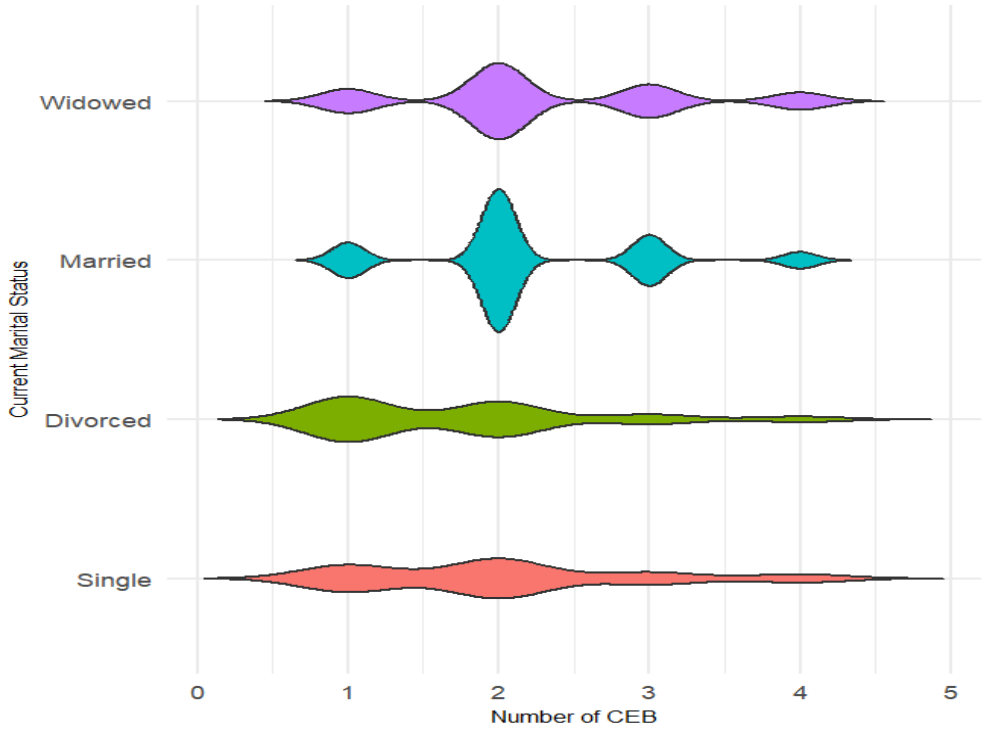

Figure 5: *Distribution of Number of CEB by Current Marital Status*

Figure 5 depicts the distribution of the number of children by current marital status. It shows the differences in fertility rates among single, married, divorced, and widowed women.

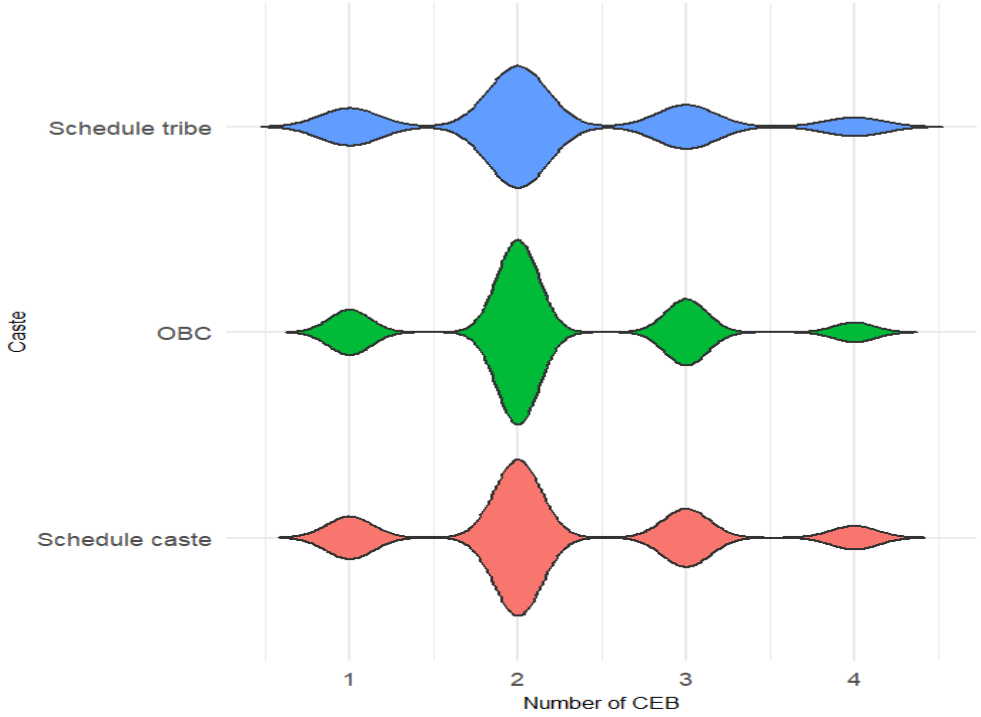

Figure 6: *Distribution of Number of CEB by Caste*

Figure 6 shows the violin plot of the distribution of the number of CEB within each caste group. Most women had two children in various caste groups, providing a clear picture.

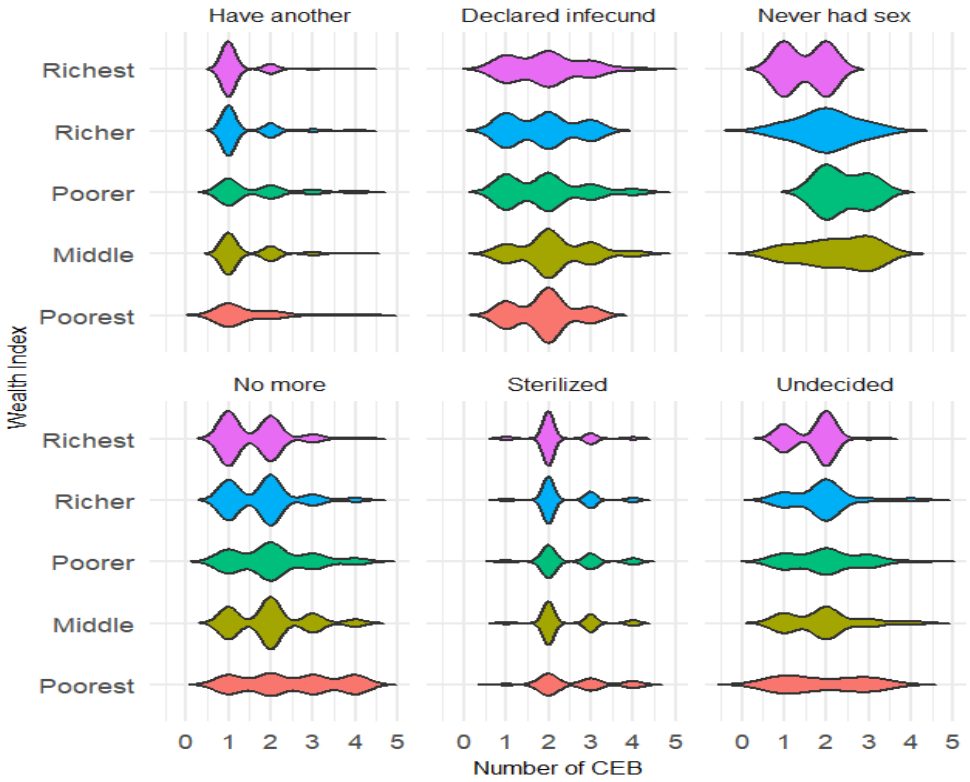

Figure 7: *Violin Plot of Number of CEB by Wealth Index and Fertility Preference*

Figure 7 visualizes the flow of the number of children across different wealth index categories and fertility preferences. It helps to understand how different wealth categories and fertility attitudes influence family size, with the flow's thickness indicating the transition's magnitude.

Table 3: *Overall model comparison by model fit characteristics*

Test Statistics	ZTP	ZTGP
Log Likelihood	-10191.22	-11520.35
AIC	20462.45	23122.7
BIC	20742.37	23409.61

Table 3 clearly shows that the loglikelihood of the ZTP model (-10191.22) is greater than that of the ZTGP model (-11520.35). According to this study, the ZTP model provides a more accurate fit for the data. The Akaike Information Criterion (AIC) of the Zero-Truncated Poisson (ZTP) model (20462.45) is lower than that of the Zero-Truncated Generalized Poisson (ZTGP) model (23122.7), indicating that the ZTP model is a better fit and has less complexity. The BIC value for the ZTP model (20742.37) is lower than the BIC value for the ZTGP model (23409.61), indicating that the ZTP model outperforms the ZTGP model in terms of BIC. BIC is known for imposing greater penalties on model complexity than AIC. Thus, compared to the ZTGP model, the ZTP model outperforms it on these criteria.

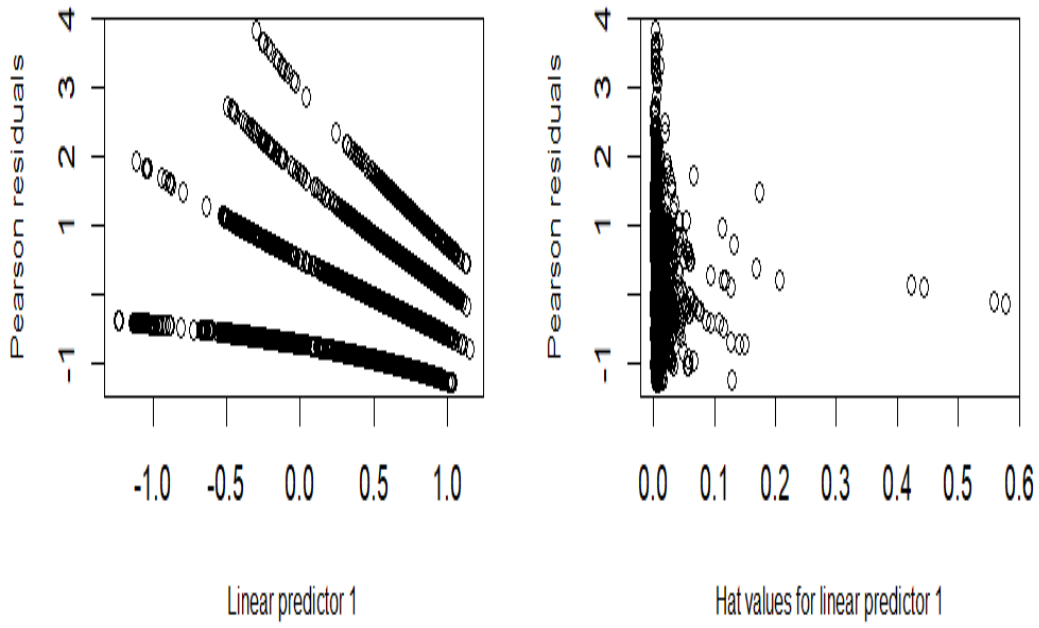

Figure 8: *Linear Predictor of ZTP model*

Table 4: *Sociodemographic, socioeconomic, and environmental factors affecting the number of children born among reproductive-aged women in AP; data from NFHS-5*

Variables	Category	N	Percentage
Place of residence	Urban	2235	27.6
	Rural	5852	72.4
Religion	Hindu	6780	83.8
	Muslim	615	7.6
	Christian	692	8.6

Type of cooking fuel	Electricity	57	0.7
	LPG	6692	82.8
	Biogas	7	0.1
	Kerosene	11	0.1
	Coal, lignite	16	0.2
	Charcoal	102	1.3
	Wood	1104	13.7
	Straw/shrubs/grass	35	0.4
	Agricultural crop	59	0.7
	Animal dung	2	0
	Other	2	0
Wealth index combined	Poorest	336	4.2
	Poorer	1556	19.2
	Middle	2633	32.6
	Richer	2380	29.4
	Richest	1182	14.6
Place of delivery	Home	3551	43.9
	Public	3925	48.5
	Private	611	7.6
Women age	15-19	93	1.1
	20-24	703	8.7
	25-29	1438	17.8
	30-34	1417	17.5
	35-39	1583	19.6
	40-44	1310	16.2
	45-50	1543	19.1
Current marital status	Single	42	0.5
	Married	7340	90.8
	Widowed	637	7.9
	Divorced	68	0.8
Fertility preference	Have another	859	10.6
	Undecided	182	2.3
	No more	651	8.0
	Sterilized	6146	76.0
	Declared infecund	228	2.8
	Never had sex	21	0.3
Caste	Schedule caste	2579	31.9
	Schedule tribe	801	9.9
	OBC	4707	58.2
Husband age	18-27	1667	20.6
	28-37	1513	18.7
	38-47	1684	20.8
	48-57	1699	21.0
	58 & above	1524	18.8
	Total	10522	100

Table 4 shows the percentage distribution of women in AP and provides an overview of the demographic and socioeconomic parameters that affect the number of CEBs. 72.4% of women live in rural areas, while 27.6% stay in urban areas. Most women, specifically 83.8%, belong to the Hinduism religion. The majority of respondents, approximately 82.8%, utilized LPG as their cooking fuel. However, 13.7% of respondents used wood, and 1.3% used charcoal. According to the wealth index, 29.4% of respondents belong to the wealthier category, while only 4.2% fall into the poorest category. Public facilities account for most deliveries, with 48.5%, followed by house deliveries at 43.9%. 19.6% of the participants fell within the age range of 35-39 years, while a close second was the group of respondents aged between 45-50 years, accounting for 19.1%.

Almost 90.8% of the marital status occurred by women who were married. Approximately 76% of the women have undergone sterilization, reflecting their lack of desire to have any more children. The percentage of respondents from Other Backward Classes (OBC) was 58.2%, while the percentage of respondents from Scheduled Tribes was 9.9%. The age distribution of husbands exhibits distinct age categories, with the largest proportion (21%) lying within the range of 48-57 years, followed by 20.8% coming within the range of 38-47 years.

Table 5: *Cross-tabulation of predictor factors with AP's (n=8087) children ever born count among women in their reproductive years*

Variables	Response	1	2	3	4
Place of residence	Urban	371 (30.9)	1353 (29.2)	367 (21.8)	144 (24.9)
	Rural	829 (69.1)	3274 (70.8)	1315 (78.2)	434 (75.1)
Religion	Hindu	1026 (85.5)	3962 (85.6)	1350 (80.3)	442 (76.5)
	Muslim	80 (6.7)	313 (6.8)	154 (9.2)	68 (11.8)
	Christian	94 (7.8)	352 (7.6)	178 (10.6)	68 (11.8)
Type of cooking fuel	Electricity	7 (0.6)	33 (0.7)	12 (0.7)	5 (0.9)
	LPG	1005 (83.8)	3913 (84.6)	1341 (79.7)	433 (74.9)
	Biogas	4 (0.3)	2 (0.0)	1 (0.1)	0 (0.0)
	Kerosene	3 (0.3)	3 (0.1)	3 (0.2)	2 (0.3)
	Coal, lignite	2 (0.2)	11 (0.2)	2 (0.1)	1 (0.2)
	Charcoal	14 (1.2)	45 (1.0)	28 (1.7)	15 (2.6)
	Wood	156 (13.0)	561 (12.1)	272 (16.2)	115 (19.9)
	Straw/shrubs/grass	2 (0.2)	21 (0.5)	8 (0.5)	4 (0.7)
	Agricultural crop	7 (0.6)	34 (0.7)	15 (0.9)	3 (0.5)

	Animal dung	0 (0.0)	2 (0.0)	0 (0.0)	0 (0.0)
	Other	0 (0.0)	2 (0.0)	0 (0.0)	0 (0.0)
Wealth index combined	Poorest	53 (4.4)	150 (3.2)	85 (5.1)	48 (8.3)
	Poorer	217 (18.1)	804 (17.4)	379 (22.5)	156 (27.0)
	Middle	344 (28.7)	1500 (32.4)	590 (35.1)	199 (34.4)
	Richer	358 (29.8)	1405 (30.4)	476 (28.3)	141 (24.4)
	Richest	228 (19.0)	768 (16.6)	152 (9.0)	34 (5.9)
Place of delivery	Home	523 (43.6)	2051 (44.3)	712 (42.3)	265 (45.8)
	Public	589 (49.1)	2227 (48.1)	838 (49.8)	271 (46.9)
	Private	88 (7.3)	349 (7.5)	132 (7.8)	42 (7.3)
Women age	15-19	70 (5.8)	22 (0.5)	1 (0.1)	0 (0.0)
	20-24	275 (22.9)	362 (7.8)	62 (3.7)	4 (0.7)
	25-29	262 (21.8)	907 (19.6)	222 (13.2)	47 (8.1)
	30-34	172 (14.3)	916 (19.8)	265 (15.8)	64 (11.1)
	35-39	153 (12.8)	1031 (22.3)	310 (18.4)	89 (15.4)
	40-44	122 (10.2)	717 (15.5)	338 (20.1)	133 (23.0)
	45-50	146 (12.2)	672 (14.5)	484 (28.8)	241 (41.7)
Current marital status	Single	13 (1.1)	19 (0.4)	6 (0.4)	4 (0.7)
	Married	1053 (87.8)	4262 (92.1)	1528 (90.8)	497 (86.0)
	Widowed	102 (8.5)	321 (6.9)	141 (8.4)	73 (12.6)
	Divorced	32 (2.7)	25 (0.5)	7 (0.4)	4 (0.7)
Fertility preference	Have another	602 (50.2)	196 (4.2)	46 (2.7)	15 (2.6)
	Undecided	55 (4.6)	98 (2.1)	23 (1.4)	6 (1.0)
	No more	226 (18.8)	291 (6.3)	94 (5.6)	40 (6.9)

	Sterilized	242 (20.2)	3930 (84.9)	1466 (87.2)	508 (87.9)
	Declared infecund	71 (5.9)	102 (2.2)	46 (2.7)	9 (1.6)
	Never had sex	4 (0.3)	10 (0.2)	7 (0.4)	0 (0.0)
Caste	Schedule caste	391 (32.6)	1436 (31.0)	535 (31.8)	217 (37.5)
	Schedule tribe	135 (11.3)	440 (9.5)	159 (9.5)	67 (11.6)
	OBC	674 (56.2)	2751 (59.5)	988 (58.7)	294 (50.9)
Husband age	18-27	251 (20.9)	966 (20.9)	337 (20.0)	113 (19.6)
	28-37	202 (16.8)	875 (18.9)	333 (19.8)	103 (17.8)
	38-47	265 (22.1)	954 (20.6)	345 (20.5)	120 (20.8)
	48-57	262 (21.8)	959 (20.7)	353 (21.0)	125 (21.6)
	58 & above	220 (18.3)	873 (18.9)	314 (18.7)	117 (20.2)

According to the data in Table 5, rural women have a greater percentage of larger families. Specifically, 78.2% of rural women have three children, while 75.1% have four or more children. Meanwhile, only 21.8% of urban women have three children, and 24.9% have four or more children. According to this study, living in rural areas is linked to greater fertility rates. Among all religious groups, Hindu women have the greatest percentage in all categories of childbearing age. Muslim women exhibit a greater proportion in the higher CEB categories, specifically 9.2% (154) for three children and 11.8% (68) for four or more children, in comparison to Christian women who had 10.6% (178) for three children and 11.8% (68) for four or more children.

Among women, 84.6% use LPG as their main cooking fuel, which affects all CEB groups. Those who use wood as fuel have greater percentages in bigger family sizes, with 16.2% having three children and 19.9% having four or more children. Women with two children have a greater proportion (22.3%) in the CEB categories for the age group of 35-39 years. Among all CEB categories, married women had the greatest percentage, specifically 92.1% (4262). The majority of women who undergo sterilization belong to the higher CEB groups, with 87.2% having three children and 87.9% having four or more children. Most women in the Other Backward Classes (OBC) in all Central Employment Bureau (CEB) categories have higher percentages. The spouses of these women are between the ages of 38-47 and 48-57. Additionally, more women with two children (954, 20.6%) and (959, 20.7%) fall into this category.

Table 6: *Results from the ZTP Model*

Variables	Category	Estimate	Std. Error	z-value	P-value	IRR	95% Wald Confidence Interval IRR	
Intercept		-1.088	0.271	-4.012	0.000***	0.337	0.198	0.573
Place of residence (Ref: Urban)	Rural	0.003	0.023	0.117	0.907@	1.003	0.958	1.050
Religion (Ref: Hindu)	Muslim	0.149	0.034	4.410	0.000***	1.161	1.087	1.241
	Christian	0.064	0.034	1.892	0.059*	1.066	0.998	1.139
Type of cooking fuel (Ref: Electricity)	LPG	-0.039	0.108	-0.364	0.716@	0.961	0.778	1.188
	Biogas	-0.619	0.481	-1.287	0.198@	0.539	0.210	1.382
	Kerosene	0.169	0.258	0.654	0.513@	1.184	0.714	1.963
	Coal, lignite	-0.134	0.241	-0.557	0.578@	0.874	0.545	1.402
	Charcoal	0.048	0.132	0.366	0.714@	1.049	0.811	1.359
	Wood	-0.001	0.111	-0.010	0.992@	0.999	0.804	1.241
	Straw/shrubs/grass	0.042	0.169	0.248	0.804@	1.043	0.749	1.452
	Agricultural crop	-0.036	0.152	-0.240	0.811@	0.964	0.716	1.298
	Animal dung	-0.140	0.658	-0.213	0.831@	0.869	0.239	3.159
	Other	-0.360	0.659	-0.546	0.585@	0.698	0.192	2.538
Wealth index combined (Ref: Poorest)	Poorer	-0.071	0.048	-1.477	0.140@	0.931	0.848	1.024
	Middle	-0.120	0.049	-2.461	0.014**	0.887	0.806	0.976
	Richer	-0.180	0.051	-3.517	0.000***	0.835	0.756	0.923
	Richest	-0.319	0.057	-5.628	0.000***	0.727	0.650	0.812
Place of delivery (Ref: Home)	Public	-0.014	0.019	-0.751	0.453@	0.986	0.949	1.024
	Private	0.003	0.036	0.072	0.943@	1.003	0.934	1.077
Women age (Ref: 15-19)	20-24	0.603	0.202	2.990	0.003***	1.827	1.231	2.711
	25-29	0.755	0.199	3.787	0.000***	2.128	1.440	3.146
	30-34	0.776	0.200	3.886	0.000***	2.173	1.469	3.214
	35-39	0.809	0.200	4.053	0.000***	2.246	1.519	3.321
	40-44	0.905	0.200	4.531	0.000***	2.472	1.671	3.657
	45-50	1.001	0.200	5.016	0.000***	2.721	1.840	4.023
Current marital status (Ref: Single)	Married	0.215	0.140	1.531	0.126@	1.240	0.942	1.633
	Widowed	0.142	0.143	0.989	0.323@	1.152	0.870	1.527
	Divorced	-0.208	0.191	-1.090	0.276@	0.812	0.559	1.181
Fertility preference (Ref: Have another)	Undecided	0.625	0.091	6.898	0.000***	1.868	1.564	2.230
	No more	0.654	0.066	9.971	0.000***	1.923	1.691	2.186
	Sterilized	0.934	0.056	16.809	0.000***	2.544	2.282	2.837
	Declared infecund	0.609	0.083	7.337	0.000***	1.838	1.562	2.163
	Never had sex	0.763	0.195	3.912	0.000***	2.144	1.463	3.141
Caste (Ref:	Schedule tribe	-0.042	0.035	-1.194	0.232@	0.959	0.896	1.027

Schedule caste)	OBC	-0.030	0.022	-1.405	0.160@	0.970	0.930	1.012
Husband age (Ref: 18-27)	28-37	0.021	0.030	0.713	0.476@	1.021	0.964	1.082
	38-47	0.017	0.029	0.575	0.565@	1.017	0.961	1.076
	48-57	0.018	0.029	0.641	0.522@	1.019	0.963	1.078
	58 & above	0.023	0.030	0.781	0.435@	1.023	0.966	1.084

***1% Level of Significant (p-value<0.01)

** 5% Level of Significant (p-value<0.05)

* 10% Level of Significant (p-value<0.1)

@ Not Significant

Table 6 displays the Zero Truncated Poisson (ZTP) model, which analyzes the factors that influence the number of Children Ever Born (CEB) among women of reproductive age. The results indicate that the Muslim faith has a favourable effect, whereas income index categories (excluding poorer women) have a negative effect. A positive relationship exists between women's age (from 20-24 to 45-50) and fertility preference. Women who used Biogas as a cooking fuel observed a 46% lower risk (IRR = 0.539, 95% CI: 0.210-1.382), whereas those who used kerosene had an 18% greater risk (IRR = 1.184, 95% CI: 0.714-1.963), in comparison to women who depended on electricity. Women from lower socioeconomic backgrounds had a 6.9% lower risk (IRR = 1.02, 95% CI: 0.97-1.024) compared to women who give birth at home.

Women aged 20-24, 25-29, 30-34, 35-39, 40-44, and 45-50 had significantly higher incidences compared to women aged 15-19, with incidence rate ratios (IRR) of 1.827 (95% CI: 1.231-2.711), 2.128 (95% CI: 1.440-3.146), 2.173 (95% CI: 1.469-3.214), 2.246 (95% CI: 1.519-3.321), 2.472 (95% CI: 1.671-3.657), and 2.721 (95% CI: 1.840-4.023), respectively. Married women had a 24% decrease in risk (IRR = 1.240, 95% CI: 0.942-1.633) compared to single women. Divorced women had an 18.8% decrease in risk (IRR = 0.812, 95% CI: 0.559-1.181) compared to single women.

Women who have had sterilization have a 1.544 times greater risk (IRR = 2.544, 95% CI: 2.282-2.837) compared to women who have given birth to another child. Women belonging to the Scheduled Caste had a 4.1% reduced risk of fertility compared to women belonging to the Scheduled Tribe caste, with an incidence rate ratio (IRR = 0.959, 95% CI: 0.896-1.027). Respondents whose husbands were aged 58 and above and 28-37 had a 2.3% (IRR = 1.023, 95% CI: 0.966-1.084) and 2.1% (IRR = 1.021, 95% CI: 0.964-1.082) higher risk of fertility compared to respondents whose husbands were aged 18-27.

Table 7: *Results from the ZTGP Model*

Variables	Category	Estimate	Std. Error	z-value	P-value	IRR	95% Wald Confidence Interval IRR	
Intercept1		0.089	0.175	0.507	0.612@	1.093	0.776	1.538
Intercept2		-30.490	8.245	-0.004	0.997@	0.000	0.000	Inf
Place of residence (Ref: Urban)	Rural	0.002	0.019	0.088	0.930@	1.002	0.965	1.040
Religion (Ref: Hindu)	Muslim	0.099	0.028	3.531	0.000***	1.104	1.045	1.167
	Christian	0.044	0.028	1.578	0.115@	1.045	0.989	1.104
Type of cooking fuel (Ref: Electricity)	LPG	-0.026	0.089	-0.288	0.773@	0.975	0.819	1.160
	Biogas	-0.330	0.315	-1.047	0.295@	0.719	0.388	1.333
	Kerosene	0.114	0.216	0.531	0.596@	1.121	0.735	1.710
	Coal, lignite	-0.088	0.193	-0.456	0.649@	0.916	0.627	1.337

	Charcoal	0.036	0.109	0.331	0.741@	1.037	0.837	1.285
	Wood	0.000	0.091	0.000	1.000@	1.000	0.836	1.196
	Straw/shrubs/grass	0.028	0.141	0.202	0.840@	1.029	0.781	1.356
	Agricultural crop	-0.024	0.124	-0.196	0.845@	0.976	0.765	1.245
	Animal dung	-0.088	0.508	-0.174	0.862@	0.915	0.338	2.480
	Other	-0.233	0.509	-0.459	0.647@	0.792	0.292	2.146
Wealth index combined (Ref: Poorest)	Poorer	-0.048	0.040	-1.199	0.231@	0.953	0.881	1.031
	Middle	-0.080	0.041	-1.970	0.049**	0.923	0.852	1.000
	Richer	-0.120	0.043	-2.822	0.005***	0.887	0.816	0.964
	Richest	-0.207	0.047	-4.450	0.000***	0.813	0.742	0.891
Place of delivery (Ref: Home)	Public	-0.010	0.016	-0.639	0.523@	0.990	0.960	1.021
	Private	0.001	0.030	0.032	0.975@	1.001	0.945	1.061
Women age (Ref: 15-19)	20-24	0.169	0.098	1.735	0.083*	1.184	0.978	1.434
	25-29	0.250	0.096	2.605	0.009**	1.284	1.064	1.550
	30-34	0.263	0.096	2.730	0.006**	1.301	1.077	1.572
	35-39	0.286	0.096	2.961	0.003***	1.330	1.101	1.607
	40-44	0.352	0.097	3.633	0.000***	1.421	1.176	1.718
	45-50	0.420	0.097	4.355	0.000***	1.523	1.260	1.840
Current marital status (Ref: Single)	Married	0.137	0.109	1.256	0.209@	1.147	0.926	1.420
	Widowed	0.088	0.112	0.784	0.433@	1.092	0.877	1.359
	Divorced	-0.107	0.142	-0.750	0.453@	0.899	0.680	1.188
Fertility preference (Ref: Have another)	Undecided	0.251	0.062	4.019	0.000***	1.285	1.137	1.452
	No more	0.268	0.042	6.420	0.000***	1.307	1.204	1.418
	Sterilized	0.446	0.033	13.607	0.000***	1.561	1.464	1.665
	Declared infecund	0.236	0.057	4.112	0.000***	1.266	1.131	1.416
	Never had sex	0.330	0.153	2.154	0.031**	1.390	1.030	1.877
Caste (Ref: Schedule caste)	Schedule tribe	-0.028	0.028	-0.982	0.326@	0.973	0.920	1.028
	OBC	-0.020	0.018	-1.116	0.264@	0.981	0.947	1.015
Husband age (Ref: 18-27)	28-37	0.014	0.024	0.574	0.566@	1.014	0.967	1.063
	38-47	0.011	0.023	0.462	0.644@	1.011	0.966	1.058
	48-57	0.012	0.023	0.517	0.605@	1.012	0.967	1.060
	58 & above	0.015	0.024	0.612	0.541@	1.015	0.968	1.064

***1% Level of Significant (p-value<0.01)

** 5% Level of Significant (p-value<0.05)

* 10% Level of Significant (p-value<0.1)

@ Not Significant

The analysis of the ZTGP model in Table 7 provides the impact of various determinants on fertility count, specifically focusing on the number of CEB. The results indicate that factors such as the positive influence of the Muslim faith, the negative impact of being richer or richest women, the positive influence of women's age ranging from 25-29 to 45-50, and the positive influence of

fertility preference categories are all significant.

The occurrence rates of CEB (Childbearing Ever) increased by 10.4% (IRR = 1.104, 95% CI:1.045-1.167) and 4.5% (IRR = 1.045, 95% CI:0.989-1.104) for women who held Muslim and Christian beliefs, respectively, compared to those who practis loglikelihooded Hinduism. Women who use biogas and other fuel sources observe a 28% decrease in the occurrence of CEB (IRR = 0.719, 95% CI:0.388-1.333) and a 21% reduction (IRR = 0.792, 95% CI:0.292-2.146) compared to women who depend on electricity. The wealthiest women exhibited a reduced rate of childbearing (IRR = 0.813, 95% CI: 0.742-0.891) compared to the poorest women.

In comparison to women aged 15-19 years, the rates of the number of CEB among women aged 20-24, 25-29, 30-34, 35-39, 40-44, and 45-50 increased by 18%, 28%, 30%, 33%, 42%, and 52% respectively. Compared to single women, divorced women had a 10% reduced incidence risk of CEB (IRR = 0.899, 95% CI: 0.680-1.188). Similarly, the incidence rates of CEB among sterilized women increased by 56% (IRR = 1.561, 95% CI:1.464-1.665) in comparison to women with different fertility preferences. The incidence rate of CEB among women who belong to the Scheduled Tribe was reduced by 3% (IRR = 0.973, 95% CI: 0.920-1.028) compared to women from the Scheduled Caste. Similarly, women from the Other Backward Classes (OBC) had a 2% decrease in the incidence rate of CEB (IRR = 0.981, 95% CI: 0.947-1.015) compared to women from the Scheduled Caste.

IV. Discussion

This study included 8087 women between the ages of 15 and 50. The majority of these women, representing 44%, had given birth to two children. This study utilized the ZTP and ZTGP regression models to examine reproductive patterns in Andhra Pradesh. The analysis focused on the number of CEBs, which was influenced by several variables, including various sociodemographic, socioeconomic, and environmental factors.

The study found that a significant proportion of women (72.4%) live in rural areas, while a smaller proportion (27.6%) live in urban areas. The generality of the rural regions is associated with higher fertility rates, as women living in rural areas tend to have bigger families due to the limited availability of family planning services and distinct socio-cultural norms. Hinduism is the most prevalent religion, with 83.8% of women identifying as Hindus. There is variation in fertility rates among different religious groups, with Muslim women having greater fertility rates in higher CEB categories compared to Hindu and Christian women. The majority of women (82.8%) utilize LPG as their primary cooking fuel, although traditional fuels such as wood (13.7%) and charcoal (1.3%) are less prevalent. The utilization of traditional fuels has been associated with higher household sizes, most likely due to socioeconomic limitations and lifestyle aspects in lower-income households.

According to the wealth index, 29.4% of respondents are classified as rich, while only 4.2% are categorized as the poorest. Women with higher wealth tend to have fewer children, indicating the negative correlation between economic status and fertility rates. Public facilities are the most common places for delivery (48.5%), followed by home deliveries (43.9%). The selection of the birth location is impacted by factors such as ease of access, cost-effectiveness, and cultural inclinations, which subsequently affect the health results of both mothers and children. The age distribution reveals that 19.6% of participants fall within the age range of 35-39 years, while 19.1% fall within the age range of 45-50 years. A majority (90.8%) of women are married, and this is strongly correlated with reproductive patterns since married women tend to have greater rates of fertility. The prevalence of sterilization is high, with 76% of women undergoing the procedure, which indicates a clear and final decision to cease childbearing. The caste division reveals that 58.2% of the participants belong to the Other Backward Classes (OBC), while 9.9% are from

scheduled tribes, thus highlighting the socio-cultural diversity within the community.

Rural women exhibit higher fertility rates, with 78.2% having three children and 75.1% having four or more children, compared to 21.8% and 24.9% for urban women. This study highlights the impact of living in rural areas on increased fertility rates. Hindu women make up the largest proportion of all categories of CEB. In contrast, Muslim women have a greater proportion in the upper CEB categories (9.2% for three children and 11.8% for four or more children) compared to Christian women. Women who use LPG have a significant presence in all CEB categories. However, those that depend on traditional fuels such as wood tend to have higher proportions of bigger family sizes, indicating that economic and social factors affect fertility. Women between the ages of 35 and 39 have more children in the CEB categories, while married women make up the majority of all CEB categories with a percentage of 92.1%. Women who have had sterilization tend to have a larger number of CEB, which indicates their previous high fertility before treatment. Women in the Other Backward Classes (OBC) exhibit greater proportions in all CEB categories, suggesting higher fertility rates within this demographic.

The ZTP regression model study identifies significant factors, including the Muslim faith, wealth index, women's age, and fertility preference. Muslim women, women with higher wealth, women in older age groups (20-50), and those who prefer sterilization have higher rates of CEB. This analysis, like the ZTP model, additionally highlights the beneficial impact of the Muslim faith and women's age on CEB. However, the data reveals a more prominent adverse effect of money, as wealthier and wealthiest women have far lower rates of childbearing. The ZTP model indicates that the Muslim faith is positively associated with higher CEB, whereas wealth index categories, except for poorer women, show a negative association. Age is an important variable, as women between 20 and 50 have significantly higher CEB than those in the 15-19 age group. Fertility desire, namely sterilization, is an additional influential factor since sterilized women tend to have a larger number of CEB.

To manage population growth to improve mother and child health outcomes, rural areas with higher fertility rates require focused family planning and reproductive health care. The significant impact of religion on reproductive patterns highlights the necessity for culturally sensitive interventions that respond to the distinct requirements and opinions of various religious communities. The negative relationship between wealth and fertility underscores the significance of economic advancement and education in lowering childbearing rates. There is a relationship between the financial status of women and the number of children they have, indicating that enhancing economic status may result in decreased fertility rates. The strong relationship between age, marital status, and fertility suggests that reproductive health programs must focus on specific age groups and marital statuses to manage fertility rates effectively.

V. Conclusion

This study aimed to examine reproductive patterns and the determinants that influence the number of children ever born (CEB) among women aged 15 to 50 in Andhra Pradesh using count data regression models based on NFHS conducted between 2019-2021. The ZTP regression model most effectively identified the key factors influencing CEB. These factors include the Muslim religion, the wealth index of the richest individuals, women aged between 20 and 50, and fertility preferences based on sterilization. Furthermore, reproductive health programs should be customized to the specific needs of different demographic groups, particularly in rural regions, such as religion, wealth, age, and marital status, to effectively manage population growth and improve the health outcomes of mothers and children. By addressing the specific determinants of fertility identified, policymakers can develop more effective strategies to encourage sustainable population growth and improve the overall health of women and children in the region.

Ethics approval and consent to participate

The Demographic and Health Surveys (DHS) Program permits the authors to download survey data. The publicly available data do not contain personal information.

Acknowledgments

The author acknowledges the DHS Program for providing access to the use of the reviewed data.

References

[1] Agresti, A. (2019). An introduction to categorical data analysis.

[2] Aldieri, L., & Vinci, C. P. (2010). An investigation of the relation between the number of children and education in Italy.

[3] Al-Balushi, M. S., Ahmed, M. S., Islam, M. M., & Khan, M. H. R. (2020). Multilevel poisson regression modeling to identify factors influencing the number of children ever born to married women in Oman. Journal of Statistics and Management Systems, 23(8), 1357-1373.

[4] Aragaw Eshetie Aguade and B. Muniswamy. 2018. Proposed Score Test for Over-dispersion Parameter in the Multilevel Negative Binomial Regression Model, Journal of Emerging Technologies and Innovative Research, Volume 5, Issue 12, pp. 709-720, ISSN-2349-5162.

[5] Bongaarts, J. (2016). Development: Slow down population growth. Nature, 530(7591), 409-412.

[6] Bongaarts, J. 1 978. A framework for analyzing the proximate determinants of fertility. Population and Development Revie1v, 4(1), 105-132.

[7] Brinker, G., & Amonker, R. (2013). Socioeconomic development and fertility trends among the states of India. International Journal of Sociology and Social Policy, 33(3/4), 229-245.

[8] Cameron, A. C., & Trivedi, P. K. (2013). Regression analysis of count data (No. 53). Cambridge university press.

[9] Chakrabarti, A., & Ghosh, J. K. (2011). AIC, BIC and recent advances in model selection. Philosophy of statistics, 583-605.

[10] Consul, P. C., & Famoye, F. (1989). The truncated generalized Poisson distribution and its estimation. Communications in Statistics-Theory and Methods, 18(10), 3635-3648.

[11] Gurmu, S. (1991). Tests for detecting overdispersion in the positive Poisson regression model. Journal of Business & Economic Statistics, 9(2), 215-222.

[12] Haque, M. E., Mallick, T. S., & Bari, W. (2022). Zero truncated Poisson model: an alternative approach for analyzing count data with excess zeros. Journal of Statistical Computation and Simulation, 92(3), 476-487.

[13] Hilbe, J. M. (2011). Negative binomial regression. Cambridge University Press.

[14] https://rchiips.org/nfhs/NFHS5_FCTS/COMPENDIUM/Andhra_Pradesh.pdf National Family Health Survey. NFHS-5: Compendium of Fact Sheet, KEY INDICATORS State and Districts of Andhra Pradesh. (Accessed August 24, 2023). (www.rchiips.org)

[15] Ibeji, J. U., Zewotir, T., North, D., & Amusa, L. (2020). Modelling fertility levels in Nigeria using Generalized Poisson regression-based approach. Scientific African, 9, e00494.

[16] IIPS, O. (2007). National Family Health Survey (NFHS-3), 2005-06: India. Vol. I. Mumbai: International Institute for Population Sciences.

[17] India: DHS, 2019-2021 – Final Report (English) https://dhsprogram.com/publications/publication-FR375-DHS-Final-Reports.cfm.

[18] International Institute for Population Sciences (IIPS) and ICF. (2021). National Family

Health Survey (NFHS-5), India, (2019–2021). Demographic and Health Surveys, 1, 1-714.

[19] Karimuzzaman, M., Moyazzem Hossain, M., & Rahman, A. (2020). Finite Mixture Modelling Approach to Identify Factors Affecting Children Ever Born for 15–49 Year old Women in Asian Country. In Statistics for Data Science and Policy Analysis (pp. 221-236). Springer Singapore.

[20] Kiser, H., & Hossain, M. A. (2018). Estimation of number of ever born children using zero truncated count model: evidence from Bangladesh Demographic and Health Survey. Health Information Science and Systems, 7(1), 3.

[21] Lavanya, M. V., & Muniswamy, B. (2024). Exploring Reproductive Patterns: A Poisson Regression Study In Andhra Pradesh. African Journal of Biomedical Research, 27(1S), 1181-1190.

[22] Lavanya, M. V., & Muniswamy, B. (2024). A Study on The Hurdle Poisson Regression Model for Reproductive Patterns on Count Data. African Journal of Biological Sciences, Vol-6(4), 1309-1322.

[23] Melese, Z. Y., & Zeleke, L. B. (2020). Factors affecting children ever born among reproductive aged women in Ethiopia; data from Edhs 2016. World, 5(3), 66-75.

[24] Muoka, A. K., Waititu, A. G., & Ngesa, O. O. (2016). Statistical models for count data.

[25] 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.

[26] Rahman, A., Hossain, Z., Rahman, M. L., & Kabir, E. (2022). Determinants of children ever born among ever-married women in Bangladesh: evidence from the Demographic and Health Survey 2017–2018. BMJ open, 12(6), e055223.

[27] Wang, W., & Famoye, F. (1997). Modeling household fertility decisions with generalized Poisson regression. Journal of population economics, 10, 273-283.

[28] Wei-hua, Z. H. A. O., Yu, F. E. N. G., & Ze-an, L. I. (2010). Zero-truncated generalized Poisson regression model and its score tests. Journal of East China Normal University (Natural Science), 2010(1), 17.

[29] Winkelmann, R., & Zimmermann, K. F. (1994). Count data models for demographic data. Mathematical Population Studies, 4(3), 205-221.

[30] Worku, G., Tadesse, G., Arega, A., & Tesfaw, D. (2022). Determinants of the number of children born in Ethiopia, evidenced from 2019 miniEDHS: Using zero-truncated count regression models.

[31] Wulu, J. T., Singh, K. P., Famoye, F., & McGwin, G. (2002). Regression analysis of count data. Journal of the Indian Society of Agricultural Statistics, 55(2), 220-231.

[32] Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression models for count data in R. Journal of statistical software, 27(8), 1-25.