Vol.4.Issue.2.2016 (April-June)

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH

A Peer Reviewed International Research Journal

http://www.bomsr.com

Email:editorbomsr@gmail.com

RESEARCH ARTICLE

ISSN:2348-0580

STUDY OF INFLUENCE OF MACROECONOMIC DETERMINANTS ON FERTILITY USING GENERALIZED POISSON REGRESSION

S.VISALAKSHI[1] , R.GEETHA[2]

[1]Assistant Professor,Department of statistics,JBAS College for women,Chennai

[2]Associate Professor,Department of Statistics, S.D.N.B.VaishnavCollege for Women,Chennai

ISSN:2348-0580

BOMSR

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH

www.bomsr.com

ABSTRACT

Children ever born (CEB) to women in their reproductive age group 15 – 49 are used as a population healthindicator of a country.So far researchers have used Poisson regression model (McCullagh, P. and Nelder, J. (1989)), (Cameron A.C.Trivedi P.K. (1998)) to studyequi-dispersed count data as a function of a set of predictor variables. The Generalized Poisson Regression (GPR) and Negative Binomial Regression (NBR) models (Consul and Famoye (1981), Winkelmann. R, Zimmermann K.F (1994)) are frequently used to study over as well as under dispersed count data.The objective of this study is to investigate and quantify the major risk factors of CEB in India taking into consideration macro-economic determinants of fertility. Thispaper presents a comparison between Poisson Regression and Generalized Poisson Regressionto model CEB in India based on Census 2011 data.Approximate tests for the dispersion and goodness of fit measures for the models are discussed.

Key words: Poisson Regression (PR), Generalized Poisson Regression (GPR), children ever born (CEB).

1. INTRODUCTION

India is the second most populous country in the world with a current population of 1.27 billion people, more than a sixth of the word population. It is likely to yield rich demographic dividend in future decades. An indicator of growth rate is the number of children ever born to women at the reproductive age group 15 – 49.In demography the word fertility is used in relation to the actual production of children or occurrence of births, especially live births. A measure Crude birth rate (CBR) is the number of live birth per women during a given period. General fertility rate is the number of babies per 1000 women in the reproductive age group.As per the data provided by SRS the estimated CBR for India in 2011 is 20.97 ranging from a high of 28.8 in Madhya Pradesh and to a low of 11.2 in Himachal Pradesh. The mean number of CEB ranges from a high of 4 in Nagaland to a low of 3 in Uttar Pradesh.There is a relationship between socio - cultural, biological, economic factors and fertility rate. Age sex structure, availability of family planning services, social and religious beliefs, female employment, economic prosperity, poverty, infant mortality rate, conflicts, typical age of marriage, and type of residence have an impact on CEB. Increasing urbanization is an important reason that changes the pattern of parity. Measures like improving the level of education and female literacy rate, awareness about early marriage and large family size to improve the situation leads to lesser number of children ever born. Higher degree of education, gross domestic product and per capita income gives the fewer number of children ever born.In recent years researchers have shown widespread interest in modeling the relationships between socio economic factors and fertility.In many empirical studies on fertility, data are obtained from demographic survey likeNHFS, (Tripathy P.K etal (2011)), BDHS (MorshedAlam, etal (2013) to model children ever born as a function of socio economic determinants of fertility. The Poisson regression model is the most popular tool for the analysis of equi – dispersed count data as a function of a set of predictor variables (Cameron and Trivedi, 1998). It is a GLM with Poisson error structure and has a wide range of applications in the field of medicine, biology, demography and econometrics. When the response variable exhibits over dispersion or under dispersion the generalized Poisson regressionmodel (GPR) is proposed by Consul and Famoye (1992) and Famoye (1993) and Negative Binomial regression are used. A number of studies have been done based on the generalized Poisson regression model by McCullagh and Nelder (1989), Breslow (1990), Famoye (1993), Winkelmann and Zimmermann (1994), Wang and Famoye(1997),CHEN (1998), and Famoye et al 2004 to deal with over dispersion in count data.An attempt is made in this study to find out the relationship between macro-economic determinants like average household size, per capita income, proportion of people living below poverty line, female literacy rate, work participation rate and age dependency ratio and fertility rate among the Indian states.Since the fertility data used in the paper is likely to exhibit over dispersion a generalized Poisson regression model is also used and comparative results are discussed.

The rest of the paper proceeds as follows: Section 2 presents the data and variables in the study along with their measure. Section 3 outlines the Poisson regression model, generalized Poisson regression model, goodness of fit test and tests for the dispersion. Section 4 presents the results and discussions. Conclusions are drawn in section 5.

2. Data and methodology

In this study, the dependent variable is the mean number of children ever born(CEB) which is estimated as a ratio of number of live births per 1000 women.The macroeconomic determinants of CEB are the average household size, per capita income (in Rupees), female literacy rate defined as the percentage of females aged 7 and above, who knows how to read and write, age dependency ratio defined as the percentage of population aged below 15 years and more than 60 years, proportion of people living below poverty line and work participation rate which is the percentage of total workers (main and marginal) to total population.The secondary data on the determinants of children ever born for the states and union territories in India which is obtained from Census 2011 and Government resources (planning and welfare) has been used in this study.

3. Generalized Poisson Regression Model(GPR)

Suppose Y_i is the count response variable that follows a generalized Poisson distribution theProbability density function of Y_i is given by (Famoye 1993, Wang and Famoye 1997):

$$f(y_i) = P(Y_i = y_i) = \left(\frac{\lambda_i}{1+\alpha\lambda_i}\right)^{y_i} \frac{(1+\alpha y_i)^{y_i-1}}{y_i!} exp\left[\frac{-\lambda_i(1+\alpha y_i)}{1+\alpha\lambda_i}\right], y_i = 0, 1, 2, \text{----}n \quad (1)$$

$\lambda_i=\lambda_i(x_i)=\exp\left(\sum_{i=1}^{p} x_{ij}\,\beta_j\right)$,$x_i=(x_{i1}=1, x_{i2}, x_{i3},$-----,$x_{ip})$ is the i^{th}row of covariate matrix X, and β= (β_1,β_2, ------------, βp) are unknown p – dimensional vector of parameters.

$$E(Y_i/x_i)=\lambda_i,\ Var(Y_i/x_i)=\lambda_i(1+\alpha\lambda_i)^2 \quad (2).$$

The parameterα is the dispersion parameter. When α =0, mean = variance is the case of equi dispersion. When α > 0 then the variance is larger than the mean and the GPR model represents count data with over dispersion.When α < 0 the variance is smaller than the mean and the GPR model represents count data with under dispersion.

The parameter **α** and β are estimated by Newton – Raphson method. The parameter **α** is also obtained using the method of estimation by equating the Pearson chi square statistic with n-p ,itsdegrees of freedom. (Breslow (1995)) given by $\sum_{i=1}^{N}\frac{(y_i-\lambda_i)^2}{\lambda_i(1+\alpha\lambda_i)^2}$ =n-p.

3.1 Goodness of fit test

Pearson chi – square test, Likelihood Ratio test, Akaike information criterion(AIC), Bayesian information criterion (BIC) are the most commonly used measures of goodness of fit in generalized linear model. The Pearson chi square statistic is given by $\sum_{i=1}^{n}\frac{(y_i-\lambda_i)^2}{var\,(y_i/x_i)}$ follows χ^2 distribution with (n-p) degrees of freedom. The second test is based on the deviance statistic D= 2(L (y_i, y) – L(λ_i y)). Where L(y_i y) and L($\lambda_{i,}$y) are log likelihood value of the estimated value under y and λ respectively. Foran adequate model D has an asymptotically χ^2(n-p). Akaike information criterion(AIC) defined as AIC= -ln L +p. Where ln L is the log likelihood value of the estimated model and p is the number of estimated parameters.

3.2 Test for dispersion

The generalized Poisson regression model reduces to the Poisson regression model when the dispersion parameter 'α' equals zero. To assess the adequacy of the GPR model over the Poisson regression model test the hypothesis.

H_0: α=0 against H_1: α≠0 (3)

The test procedure is to use the asymptotically normal Wald type "t"statistic defined as the ratio of the estimate of α to its standard error.. If the't'statistic falls outside (- 1.96, 1.96) interval reject H_0. An alternative test null hypothesis of $\alpha = 0$ is to use the likelihood ratio statistic which is approximately Chi Square distribution with one degree of freedom.

4.Results and Discussions

Table 1 gives a summary of the respective states occupying the top and bottom positions for the variables considered in this study

Table 1: Summary Statistics.

Variables	**Top State**	**Bottom State**
Mean house hold size.	Uttar Pradesh(6)	Tamil Nadu(4)
Per capita income.	Delhi(Rs.1,73,686)	Bihar (Rs.22,890)
Female literacy rate.	Kerala (92.1%)	Bihar (51.5%)
Age dependency ratio.	Bihar (85.54)	Daman and Diu (35.14)
Work Participation rate.	Himachal Pradesh (51.9)	Lakshadweep (29.1)
Below Poverty line	Chhattisgarh (29.93)	Chandigarh (0)
CEB	Nagaland(4)	Uttar Pradesh(3)

Table 2 presents the results of parameter estimators along with the standard error of a Poisson regression model with six predictors that effect CEB in India.

Table 2: Parameter Estimates of Poisson regression.

Variables	Poisson regression		
	Estimate	**Standard error**	**Sig. value**
(Intercept)	1.744	.0078	.000
Mean household size (Mhhs)	.190	.0008	.000
Per capita income (Pcin)	-4.676E-7	1.2119E-8	.000
Female literacy rate (Flr)	-.006	4.5586E-5	.000
Age dependency ratio (Adr)	.006	5.4854E-5	.000
Below poverty line (Bpl)	.002	2.8819E-5	.000
Work Participation rate (Wpr)	.009	6.8276E-5	.000
(Scale)	1		

The fitted Poisson regression model is given as follows:
log(live births/mid year population)= (1.744+.190 Mhhs-.00000046*Pcin -.006*Flr+.006*Adr+.002*Bpl+.009*Wpr.

The estimated value of the female literacy rate and per capita income are negative and are significant which shows compare to uneducated people, educated people have a less number of children ever born. Higher the per capita income of a state less is the number of children ever born. However,we findthattheestimated mean of the outcome variable is less than its variance in case of over dispersion in the data. So we attempt to fit a Generalized Poisson Regression(GPR) which is more appropriate in such cases.

Table (3) gives the parameter estimate of the GPR model. The parameter estimate from both Poisson regression and Generalized Poisson regression models are quite similar though the standard errors are different. An estimate of the dispersion parameterα=5497.623.

Table 3: Parameter Estimation of Generalized Poisson Regression.

Variables	**Generalized Poisson Regression**.		
	Estimate.	**Standard error.**	**Sig. value.**
(Intercept)	1.744	.5748	.002
Mean household size.	.190	.0574	.001
Per capita income.	-4.676E-7	8.9856E-7	.603
Female literacy rate.	-.006	.0034	.102
Age dependency ratio.	006	.0041	.171
Below poverty line	.002	.0021	.259
Work Participation rate.	.009	.0051	.063
(Scale)	5497.623		

Table (4) gives the goodness of fit of the Generalized Poisson Regression model.

Table 4: Goodness of fit

Types of Methods	Value	Df	Value/df
Deviance	157463.800	28	5623.707
Pearson Chi-Square	153933.447	28	5497.623
Log Likelihood	-78974.710		

Table 5 represents the likelihood ratio test for the overall fit of the model and this test is significant.

Table 5: Test for overall Fit.

Likelihood Ratio Chi-Square	Df	Sig.
212.425	6	0

Table 6 represents the tests of model effects. The mean house hold size is significant compared to all other explanatory variables. The estimate of dispersion parameter using GPR model is positive indicating over dispersion. The standard error is estimated as 542.23 and α=5497, since p value is <.05 the dispersion parameter is significantly different from zero. Thus we conclude that GPR is more adequate to model CEB as a function of macro-economic determinants of fertility.

Table 6: Tests of Model Effects.

Source	Type III		
	Wald Chi-Square	**df**	**Sig.**
(Intercept)	9.206	1	0.002
Mean hh size	10.994	1	0.001
Per capita income (Rs.)	0.271	1	0.603
Female literacy rate	2.67	1	0.102
Age dependency ratio	1.872	1	0.171
BPL	1.275	1	0.259
Work participation rate	3.457	1	0.063

5.Conclusion

Our study demonstrates that the GPR model is an appropriate model to identify predictors affecting children ever born in the states of India for anover dispersed data. Mean house hold size contributed significantly to CEB. Female literacy rate and per capita income have an Inverse relationship with children ever born while other factors like work participation rate, population density , below poverty line and age dependency ratio have a positive association with children ever born based on the data from census 2011.

References

[1]. McCullagh, P. and Nelder, J. (1989). Generalized Linear Models. 2nd ed., Chapman and Hall, London.

[2]. CHEN, Q.G., 1988: The analysis of life table and follow – up data with co variables using Poisson regression model. Biom. J.30, 351 – 358.

[3]. Consul P. C FamoyeF "Generalized Poisson Regression Model", Communication Statistics (Theory and Methodology), 1992, Vol.2, No.1, 89 – 109.

[4]. Famoye F. Restricted generalized poison regression mode, Communication Statistics (Theory and Methodology), 22 (1993) 1335 – 1354.

[5]. NFHS – 1 1992 – 93. Indian Institute of Population Sciences, Mumbai, 1995.

[6]. Wang, W. and Famoye, F. (1997). Modeling household fertility with generalized Poisson regression, Journal of Population Economics, 10, 1 – 11.

[7]. Wang W. Famoye F., "Modeling Household Fertility Decisions with Generalized Poisson Regression", Journal of Population Economics, 1997, Vol.10, 273-283.

[8]. Cameron A.C. , Trivedi P.K., Regression Analysis of count data, Cambridge University press, New York, 1998.

[9]. NFHS –II 1998 – 99. Indian Institute of Population Sciences, Mumbai, 2000.

[10]. Tripathy P.K and Sarangi P.K., Proximate determinants of fertility in India, The journal of Family Welfare Vol.50, No, December 2004.

[11]. MorshedAla metal., predictors of the number of under – five malnourished children in Bangladesh: application of the generalized Poisson Regression model. BMC Public Health 2013.

[12]. Winkelmann R., Zimmermann K.F (1994), Count Data models for demographic data. Mathematical population Studies 4: 205 – 221.

[13]. Consul P.C, Jain G.C: A generalization of the Poisson distributions. Technimetrics 1973, 15(4)791 – 799.