

Addressing Corner Solution Effect for Child Mortality Status Measure: An Application of Tobit Model

Hafiz M. Muddasar Jamil Shera & Irum Sajjad Dar

College of Statistics and Actuarial Sciences, University of the Punjab

Email: mudassirjamil775@gmail.com

DOI: 10.6007/IJARBSS/v4-i12/1340 URL: <http://dx.doi.org/10.6007/IJARBSS/v4-i12/1340>

Abstract

Although various models regarding corner solution data have been suggested, it is known that corner solution data is too complicated to predict the model. A corner solution is a unique solution in the sense that it is characterized by a mass-point at zero and a long right tail. For the present study, the data was taken from MICS collected by Punjab Bureau of Statistics (BOS) during 2011-2012. However, this study is based on the particular data. The child mortality within five-year interval was taken as a response variable. In the data obtained, 65% of the response variable observed to be zero. Tobit model reveals that it is the most accurate and widely used technique in this regard. The implementation of Tobit model shows that the significant predictor of child mortality in a family within five year are women age, postnatal checkups, antenatal care, total children ever born, and wealth index score. Since the majority (i.e. 74.4%) of mother age in the data is between 20 to 35 years. The results indicate that the chances of child survival are relatively higher for the mothers of age group of 20-35 years. Furthermore, Child mortality is highest for the families with low wealth index score. In addition, child mortality percentage tends to decrease with the improvement in postnatal and antenatal care facilities. The specification of Tobit model also used to verify the result of the Tobit model.

Key words: *Child Mortality, Corner Solution, Tobit Model, Normality, Homogeneity.*

1. Introduction

One of the overarching goals of World Fit for Children (WFFC) and the Millennium Development Goals (MDGs) is the reduction of child mortality under-five year by two-thirds in between 1990 to 2015 (UNICEF, 2006). In order to achieve this goal, monitoring progress and efforts are important but they are identifying difficult cost effective strategies. In addition, more resources should be advocated to be directed to child health sector by the many International agencies.

Child mortality within five year is one of the sensitive indicator of the socioeconomic and health status of a community, because it depends on the socioeconomic conditions of environment (Madise et al 2003). The Punjab's (province of Pakistan) under-five mortality rate (U5MR) was estimated at 111 deaths per 1,000 births. This means that around one in nine children born in the Punjab die before reaching their fifth birthday. These figures show no improvement since

the last five years (Punjab MICS 2007-2008). The main purpose of present study is to evaluate the most appropriate technique for the corner solution data to evaluate the relative importance of biosocial and economic factors that are associated with child mortality in Punjab, Pakistan.

A corner solution is a unique solution in the sense that it is characterized by a mass-point at zero and a long right tail. However, it is continuously distributed above zero. (I.e. children ever born, families, individual behaviors, or firms are being modeled). For several nontrivial division of the population, these types of behaviors are known as corner solutions (Amemiya, 1998). Household expenditures of different categories of durable goods are zero when household level of income is equally distributed (Greene, 2003). If some constraints and many zeros occur in the range of the explained variable, it is very difficult to estimate the model of simple linear regression. In this scenario, we mostly used logistic regression and convert response variable as binary outcome (0, 1). However, we lost much information regarding response variable (child mortality under-five). The technique that is mostly used in this scenario is Tobit model (Goldberger, 1964). Furthermore, Tobit model is used when our focus on the observation of restricted values and check the relationship between criterion variable and the other predictors (Wooldridge, 2002).

2. Literature Review

According to Austin et al., (2000), health status measurement produced censoring (floor and ceiling) effect and researchers wanted to investigate the relationship between measure of health status (MHS) and determinants of health. In health status measurement, censoring point formed to biased coefficient estimates. In this situation, Tobit model showed the most appropriate tool to handle censored variable in health related problems. The conditional distribution of the MHS had identity (same) variance, so estimates of Tobit model have superior performance as compared to the estimates of multiple linear regression. Many researchers used the Probit and logit models to handle health status score. These models gave inconsistent results and they were difficult to interpret the actual health score accurately. Therefore, Tobit model is more appropriate tool than the other ones for the type of health related problems. Fair (1978) developed a model that explained the individual's time allocation among work and two other natures of leisure time activities; time spent with paramour, and time spent with children. The Fair extracted a sample of 602 observations from the two recent magazines (1978). Criterion variable is the responses of extramarital affairs of men and women that presently married for the first time. Finally, there were preponderance of zeros in the criteria variable, OLS gave biased, and inconsistent results. Moreover, two basic assumptions of normality and homoscedasticity failed with the OLS estimation. Fair used the estimation technique for the structure of this study is called Tobit model. The estimation procedure of Tobit model achieves unbiased, consistent estimates as contrast with OLS regression, and disturbance term follows required normality assumption.

3. Material and Methods

In this study we have focused on the application of the Tobit model on child mortality within five year of age interval data for corner solution effects. Tobit Model is also an extension of Probit model (James Tobin, 1958). Tobit model is one important category to deal both continuous and discrete parts. It is very popular method to dealing the censored as well as corner solution data. Different graphical methods are used to verify the assumptions of Tobit model. In addition, different goodness of fit criteria including Pseudo R square, likelihood ratio test (LR), Akaike`s information criterion (AIC) and Bayesian information criterion (BIC) is used to assess the models` performance.. The specification of Tobit model also used to verify the result of the Tobit model.

4. Data sources for our study

In this study, child mortality data of one dependent and seven other independent variables were selected from the MICS (2007-2008). Contribution of response variable for our study is defined as child mortality under-five year interval. Response variable can be seen as corner solution data and coded as '0' stands for no dead kid in a family. In addition, "1" stands for children who died between the age group from 0 to 6 months (hour, day, week) and '2' stands for the children who died between the age group of 7 months to 12 months. In the same vein, the next age groups of the children will be coded till five year (i.e. 60 months). Seven explanatory variables have been included in this study and determine how children mortality under-five is affected by these independent variables. i.e. Education of household, women age, wealth index score, distance to the nearest health facility (in minutes), total children ever born, antenatal and postnatal care. The variables of antenatal care and postnatal care have been assigned dummy variables named as approved source (doctor, woman health visitor, nurse/midwife) coded as '1' and non-approved source (traditional birth, relation/friend, other) coded as '2'.

5. Estimation Results

In this section, Tobit model technique was to evaluate the overall effects of seven independent variables on the explained variable i.e. child mortality under-five. All results including significance level and coefficients are show in the following table.

Table 1.1: Tobit Model estimation

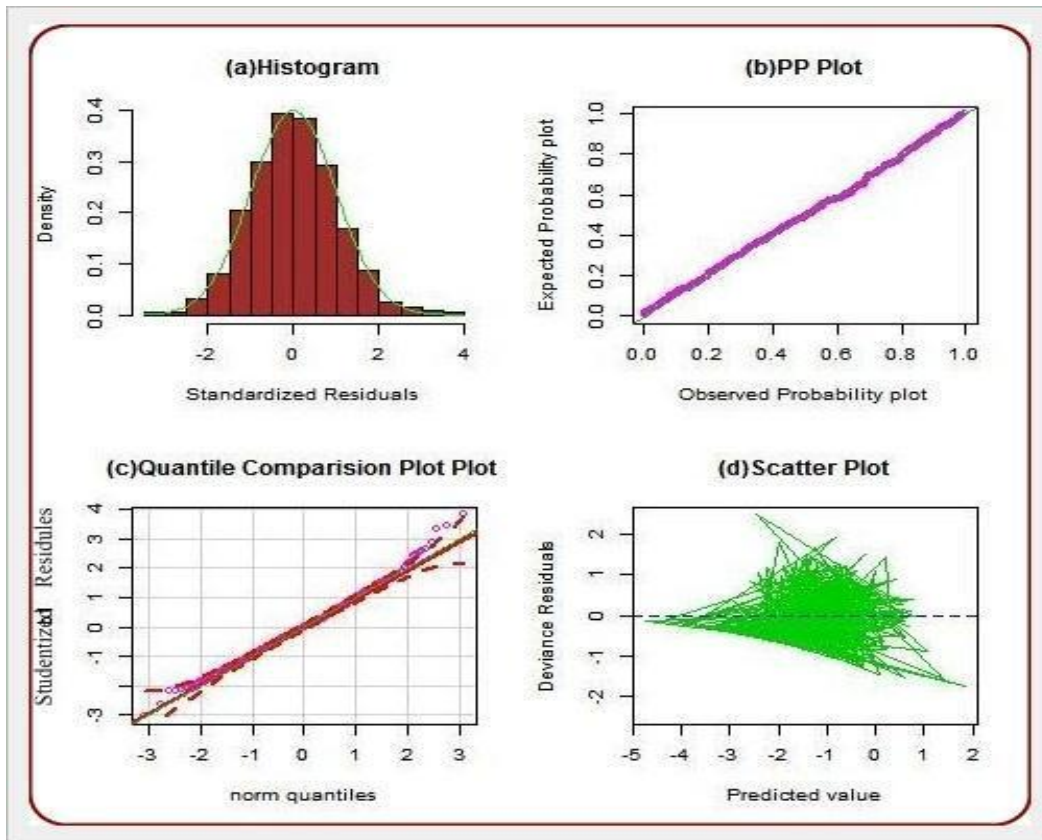
Variables	Tobit Model	
	Un-standardized Coefficients (Std. Error)	P Value
Intercept	-2.70259 (0.25763)	0.0000
Z1	-0.06413 (0.05338)	0.22954
Z2	-0.21031 (0.05921)	0.00038
Z3	-0.36102 (0.08395)	0.0000
Z4	0.10005 (0.07535)	0.18425
Z5	0.54221 (0.03067)	0.0000
Z6	0.26855 (0.14591)	0.06569
Z7	0.28854 (0.16019)	0.07165

After fitting the Tobit model on corner solution data, we can easily see that among the seven independent variables, five variables Z2 (Women age), Z3 (Wealth index score) and Z5 (Total children ever born) Z6 (Antenatal Care), and Z7 (Postnatal Checkups) indicates significant behavior at 10% level of significance, and on the other hand; the variables Z1 (Education of household head) and Z4 (Distance to the nearest health facility (in minutes)) indicates insignificant behavior in Tobit model.

5.1. Model Diagnostic Checks

Before going to further meaningful conclusions about the model analysis, we have to check the validity of model. Therefore, we check whether the above Tobit model fulfills the assumptions for predicting the criterion variable. In this section, we check important model assumptions (normality and homogeneity) by using different graphical methods. These model assumptions verify whether model is appropriate or not for the data and it can be used for the prediction purpose or not. Scatter plot is developed to assess homogeneity of variances. PP Plot, Quartile comparison plot, and histogram are used for checking the normality. Figure (1) shows that histogram, PP plot, quartile comparison plot, and scatter plot.

Figure 1: Normality and Homogeneity Graph of Tobit Model



In the above figure, histogram, PP plot, and QQ plot of Tobit model fulfill the normality assumption. We plot the Standardized residuals against the predicted value for Tobit model. On the other hand, Tobit regression model suggest homogeneity of variances. Finally, Tobit model is better approach to fulfill the assumption of normality and homoscedasticity when the data is in form of corner solution.

5.2. Specification of Tobit Model

We estimate the probit model to assess the specification of the model. It is because of the observation that both models are same with respect to structure but different with respect to measurements. Tobit model provides more information, so these estimates of the β are more consistent. As Greene (2003, pp. 776) points out that the probit estimates should be consistent for $\beta_{\text{tobit}}/\sigma_{\text{tobit}}$. Nevertheless, these results are done only if Tobit model is correct. In the below results we can easily examine the Tobit model is approximately consistent.

Table 1.2: Specification of Tobit Model (for corner solution)

Variables	β_{tobit}	β_{probit}	$\beta_{\text{tobit}}/\sigma_{\text{tobit}}$
Intercept	-2.70259	-1.71551	-1.485756
Z1	-0.06413	-0.03566	-0.03525
Z2	-0.21031	-0.11190	-0.11561
Z3	-0.36102	-0.19873	-0.19847
Z4	0.10005	0.05521	0.05502
Z5	0.54221	0.29873	0.29808
Z6	0.26855	0.14624	0.14763
Z7	0.28854	0.15892	0.15862
σ_{tobit}	1.819		

6. Conclusion and Discussion

The main purpose of this study is to evaluate the most appropriate technique for the corner solution data of child mortality under-five i.e. many observations are as zeros in dependent variable. As the above results shows that Tobit model is a more appropriate model for such data and explains the true relationship between dependent variable and the significant factors. By employing Tobit model, it is observed that child mortality within five year influenced by six factors namely i.e. women age, postnatal care, antenatal care, total children born, and wealth index score, distance to the nearest health facility (in minutes), and education of household head. The assumption of Tobit has been also conducted through the graphical and different model selection criteria. If the conditional distribution has uniform variances, Tobit model has superior performance with respect to estimates, normality, and homogeneity. The mother's age is an important factor in the determination of survival of child. The chances of child survival are higher for the mothers in an age bracket of 20 to 35 years (Galway et al. 1987; Kaldewei & Pitterle, 2011). In the same vein, Mondal, Hossain & Ali (2009) found that the child mortality rate is relatively lower for the mothers of age group of 20-29 years. The results of present research are also in accordance with these findings. Since a high majority of our respondents (i.e. 74.4%) are in an age group between 20 to 35 years. Therefore, our dataset contains a very high ratio no dead kids which is also demonstrated in the form of large proportion of zeros (i.e. 65.5%).

Wealth index score describes income level of the responding families. The results show that wealth index score variable is related with the child mortality significantly. The high wealth index score indicates high income level of the families. If a family has a high income level, it would be able to provide better antenatal and postnatal care to the baby as compared to family with low income. Therefore, the chances of child mortality are decreased with the increase in income level of the families. Our results show that the child mortality is highest for the families with lowest wealth index score. These results are in accordance with the research findings of Mustafa & Odimegwu (2008).

Moreover, the antenatal and postnatal cares are also very important factors in determining the survival chances of the children. Mondal, Hossain & Ali (2009) found that a better natal care can ensure low rate of child mortality. The percentage of child mortality levels are lower with visitation of health worker than none visitation of health worker. The results of present research also show that the child mortality percentage tends to decrease with the improvement in natal care facilities.

References

- Amemiya, T. (1998). *Advance Econometric*(2nd Ed.). United states of America: Library of congress publication data.
- Austin, C. P., and Escobar, M. (2000). The use of the tobit model for analysing measures of health statues. *Quality of life insurance* , 5(1) ,901-910.
- Dobson, A. J. (2002). *An Introduction to Generalized Linear Models* (2nd Ed.). Washington: Chapman & Hall.
- Fair, R. C. (1978). A Theory of Extramarital Affairs. *Journal of Political Economy* 86(1), 45-61.
- Greene, W. H. (2003). *Econometric Analysis* (5th Ed.). New York University: Prentice hall.
- Goldberg, S. A. (1964). *Econometric Theory* (1st Ed.). John Wiley & Sons.
- Kutner, H. M., Nachtsheim, J.C., and Neter, J. (2004). *Applied Linear Regression Models*(4th Ed.). Boston: Mc Graw Hill.
- Kaldewei, C., and Pitterle, I. (2011). Behavioural Factors as Emerging Main Determinants of Child Mortality in Middle-Income Countries: A Case Study of Jordan. *Working Paper*.
- Madise, N.J (2003): Infant mortality in Zambia: Socioeconomic and demographic correlates. *Social Biology*.

Montgomery, C. D., and Peck, A. E. (2001). Introduction to linear Regression analysis. New York: John Wiley & Sons, Inc.

Mondal, N. I., and Ali, K. M. (2009). Factors Influencing Infant and Child Mortality: A Case Study of Rajshahi District, Bangladesh. *J Hum Ecol*, 26(1), 31-39.

Mustafa, E. H., and Odimegwu. (2008). Socioeconomic Determinants of Infant Mortality in Kenya: Analysis of Kenya DHS. *Humanities and Social sciences*, 2(2).

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, 26, 24-36.

Wooldridge, J. M., and Schmidt, P. (1999). Efficient estimation of Panel data Models with strictly Exogenous Explanatory variables. *Journal of Econometric*, 93, 177-20.