RESEARCH REVIEW JOURNALS
Volume-02
Issue-09
September-2017

ISSN: 2455-3085 (Online)
RESEARCH REVIEW International Journal of Multidisciplinary
www.rrjournals.com

An Application of Tobit Regression Analysis to Dairy Sector for Gujarat State

[1]Samir B. Pandya , [2]Dr. C. D. Bhavsar

[1]*Research Scholar, Statistics, A. A. Patel Commerce College, Gandhinagar, Gujarat (India)*
[2]*Research Guide, Professor, Department of Statistics, Gujarat University, Ahmedabad, Gujarat (India)*

ARTICLE DETAILS

Article History
Received: 05 August 2017
Accepted: 10 Sep 2017
Published Online: 15 Sep 2017

Keywords
Tobit Regression, Multiple Regressions

ABSTRACT

Multivariate objects when interrelated have been pressed in two forms either it can be timed stack or cross sectional. The multidimensional studies can be evaluated in third form that compiles as binary terms. As an when the required information can classify with their nature and the multivariate computations can predict confusing results. It can represent in terms of specific code these coding can be operated for multivariate tasks that caused to binary regression. This research is carried out for the dairy sector data for listed 26 district of Gujarat states for 13 various parameters. The parameters are defined based on their median. Higher the value of parameter defined as 1 and lower the median is determined as 0; than the application of OLS has been applied in terms of tobit regression. The results are derived for binary outcomes which can be extended for future estimations.

1. INTRODUCTION TO TOBIT REGRESSION

The Tobit regression is a factual model proposed by J. Tobit in 1957, utilized so as to clarify the connection between a non-negative ward variable and an autonomous one expecting that there is a dormant variable which directly relies upon the free one trough a parameter (B) that decides the connection between the free and inactive factors. As per Varbak (2005) tobit relapse is typically the best model when the reliant variable is persistent and has a compelled fury, speaks to a positive variable. This factual model has recently been utilized in concentrates, e.g., Trefalsi et al. (2009), Hessian and Nejear (2010), yet in addition the pseudo R^2 can be registered by Effren, Fedan, Colen and Snael R Software likewise gives pseudo R^2 dependent on Feddan's recipe, as in this study while applying the Tobit validate measures.

The construction of basic tobit model can be represented as follows:

$$\widehat{Y}_i = b^a X_i + \in_i, \qquad here\ i = 1, 2 \ldots m$$

$$Y_i = \widehat{Y}_i = \begin{bmatrix} Y_i = \widehat{Y}_i, & if\ \widehat{Y}_i > 0 \\ Y_i = 0, & if\ \widehat{Y}_i \leq 0 \end{bmatrix}$$

It is defined here, $\in_i$ = random error of the model, it is shows the unobserved variables are affected the exogenous variable$\widehat{Y}_i$. The $\widehat{Y}_i$ is measure for normally affected by independently and identical distributed. In construction of model, $\widehat{Y}_i$ shows the hidden impact of variables and it can be obtained by using linear regression and follows the condition of $Y_i = \widehat{Y}_i,\ if\ \widehat{Y}_i < 0$. On other hand Y_i and X_i represents the independent variables having i = 1, 2 ….m values. Thus, in simple terms it can be represented as:

$$Y_i = \begin{bmatrix} b^a X_i + \in_i, \\ 0, \quad otherwise \end{bmatrix}$$

While (Hessian, 2003) has restructured Tobit model and supporsed the dependent parameters observes Y_i for given i = 1, 2 ….m values. The revised model was represented as:

$Y_i = maximum\left(\widehat{Y}_i, 0\right)$, It shows the estimated value of latent variables and the errors are estimated as: $\in_i \sim N(0, \sigma^2)$, Thus, $\frac{\widehat{Y}_i}{X_i} \sim N(b^a X_i,\ \sigma^2)$

$$V_i = 1, \begin{bmatrix} Y_i = \widehat{Y}_i, & if\ \widehat{Y}_i > 0 \\ Y_i = 0, & if\ \widehat{Y}_i \leq 0 \end{bmatrix}$$

Due to construction of V_i the ML (Maximum Likelihood) function can be constructed as:

$$ML_i = \prod_{i=1}^{m} P(\hat{Y} < \omega)^{1-V_i}$$

and,

$$\prod_{i=1}^{m} [P(\hat{Y} < \omega).\left(\frac{\hat{Y}}{\hat{Y} \geq \omega}\right)^{V_i}]$$

Thus the stated function can be rewrite as:

$$P(\hat{Y} < \omega) = P(b^a X_i + \in_i) < \omega$$

The construction can be divided by variance of error can be:

$$= P\left(\frac{b^a X_i + \in_i}{\sigma_\epsilon^2} < \frac{\omega}{\sigma_\epsilon^2}\right)$$

$$= P\left(\frac{\varepsilon_i}{\sigma_\epsilon^2} < \frac{\omega - b^a X_i + \in_i}{\sigma_\epsilon^2}\right) < \varphi(\frac{b^a X_i}{\sigma_\epsilon^2})$$

When, $\omega = 0$, that:

$$P(\hat{Y} < \omega) = \varphi\left(\frac{-b^a X_i}{\sigma_\epsilon^2}\right) = 1 - \varphi\left(\frac{b^a X_i}{\sigma_\epsilon^2}\right)$$

$$P(\hat{Y} \geq \omega) = 1 - [\frac{\omega - b^a X_i}{\sigma_\epsilon^2}]$$

When $\omega = 0$, than

$$P(\hat{Y} > \omega) = \varphi\left(\frac{b^a X_i}{\sigma_\epsilon^2}\right)$$

Also,

$$\therefore f[\frac{\hat{Y}}{Y} > \omega] = \frac{\frac{1}{\sigma}\varphi\{(\hat{Y} - \frac{b^a X_i}{\sigma_\epsilon^2})\}}{P(\hat{Y} > \omega)}$$

$$= \frac{\frac{1}{\sigma}\varphi\{(\hat{Y} - \frac{b^a X_i}{\sigma_\epsilon^2})\}}{1 - \varphi\{\frac{(\hat{Y} - b^a X_i)}{\sigma_\epsilon^2}\}}$$

The above calculation supports the maximum likelihood function as;

$$ML = \prod_{i=0}^{m}[1 - \varphi(\frac{b^a X_i}{\sigma_\epsilon^2})].\prod_{i=0}^{m}(\sigma_\epsilon^2)^{-1}.\varphi\,(\frac{\hat{Y} - b^a X_i}{\sigma_\epsilon^2})$$

All $\varphi(.)$ computed in functions are representing probability density functions and are tested for normality. The base tobit regression can be supported by the above construction of functions.

If the values of Y_i and X_i are not observed when, $\hat{Y} \geq \omega$. The results of dairy sector data can be judged for the final model construction of tobit regression analysis.

2. APPLICATION TO DAIRY SECTOR DATA OF GUJARAT:

An experimental construction of tobit model can be represented for 26 districts of Gujarat state for selected 7 variables (sorted based on the step regression analysis) is represented as:

$$M_i = \varphi + \omega_1 X_1 + \omega_2 X_2 + \cdots + \omega_7 X_7 + \epsilon_i$$

Here, M represents total number of states under study for dairy sector of Gujarat state; X_1 to X_i are presented as: X_1 is presenting total bovine population, X_2 is presenting total female animal population, X_3 is coded for bread animal population, X_4 shows milch animals, X_5 as sex dry ratio in animals, X_6 presents wet dry ratio in animals and X_7 is coded for vaccination institutes. The last term added to model is presenting an error term of tobit model. The model leads to φ *and* ω for determining their estimated values. The constructed model is follows the condition of normal distribution. The data are coded to binary form and are presented for appropriate calculation of each of the parameter studied for dairy sector for interpreting the results. All seven selected parameters of dairy sector are computed and presented in table 1.1.

Table 1.1: Tobit, using observations 1-650

Dependent variable: MP

	Coeff.	*SE*	*z*	*p-value*	
Const.	−0.532	0.0739	−7.202	<0.0001	***
TB_POP	0.217	0.115	1.875	0.0609	*
FA_POP	0.479	0.199	2.406	0.0161	**
BR_POP	−0.203	0.181	−1.118	0.2634	
IMA_POP	0.521	0.151	3.440	0.0006	***
MI_A	0.153	0.221	0.6900	0.4902	
MA	0.0013	0.0482	0.02750	0.9781	
SR	−0.159	0.0506	−3.157	0.0016	***
WDR	−0.042	0.0479	−0.8798	0.3790	

LS	−0.149	0.0585	−2.562	0.0104	**
PC_MA	0.395	0.0626	6.311	<0.0001	***
VI	0.0405	0.0629	0.6442	0.5194	
VA	0.187	0.0689	2.714	0.0066	***

Chi-square(12)	1063.543	p-value	4.1e-220
Log-likelihood	−382.7523	AIC	793.5047
BIC	856.1823	Hannan-Quinn	817.8158

The basics of tobit regression analysis is presented in table 1.1. The calculation is made based on the shortlisted variables for colliniarity. The lists of final seven variables are presented with construction of tobit model. The model summery shows that the standard value of error is found 0.506767 (0.0253646) with least p-value (2.28502e-019). The value of standard error is compiled for left-censored observations: 327 and Right-censored observations: 0. It shows the normality of data is negatively skewed towards the milk production. It is necessary to test the normality of residual with reference to testing null hypothesis that error is normally distributed. The results of null hypothesis are tested by taking chi-square test and p-value of the model. The result of chi-square statistics is found 85.84 with p-value 4.1e-220. The p-value indicates higher level of acceptance of null hypothesis. On other hand the chi-square value is computed 1063.54, which is higher than the significance value thus; the statistical null hypothesis is rejected in terms of dairy sector data for concluding the errors. Table 1.1 presented the construction of model which is now shows final format of construction of estimated $\varphi = -0.53\ and\ all\ \omega_i$ values are finally presented in form of function as follows:

$$MP = -0.53 + 0.217\,TB\_POP + 0.479\,FA\_POP - 0.203\,BR\_POP + 0.521\,IMA\_POP + 0.153MI\_A + 0.0013\,MA - 0.159\,SR - 0.042\,WDR - 0.149\,LS + 0.395\,PC\_MA + 0.040\,VI + 0.187\,VA$$

The Tobit model is runs for testing the sigma value and error value in accordance to all selected twelve parameters. The final model is presented as above. It is observed that the slope of model is negative towards total production of milk in Gujarat. The impact of each of the parameters are tested individually, the dairy sector is mainly have effect due to the listed variables. The result of regression model shows that the total bovine populations of animals have positive impact on model. It shows 21.7% impact. The proportions of female animal population have 47.9% positive impact on model. The In milch animals have highest impact 52.1% on model. The milch animals, Male Animals, per capita milk availability, veterinary institutions and vaccination have positive impact on model. It is computed 15.3%, 0.13%, 39.5%, 4% and 18.7% respectively. Similarly, the breed able animal, sex ratio, wet dry ratios and Livestock have negative impact on model. They are recorded as 20.3%, 15.9%, 4.2% and 14.9% respectively. The positive impact of variables on negative slope shows cross relation to the model and negative values of the parameters have positive relation to the model.

It is concluded that female animal population, in milch animal population and wet dry ratio in animals should be improved for healthier production of total milk productions. It is also observed that the vaccination institutes are developed for taking care of animal by government is having highest effect, which should be managed properly for their best outcomes.

The results of model statistics are also justify the best fit results to the model. The values of testing the model are Log-likelihood, Schwarz criterion (BIC), Akaike criterion (AIC) and Hannan-Quinn. All said tests have least values compare to other tobit constructions. It is also noticeable here the error term of the model is computed least. The error model is tested at 5% level of significance for goodness of fit. Chi-square value for hypothesis that the error of model is not normally distributed is rejected towards the computed value 1063.543. It shows that the model has computed for least value or lower error for residuals. The p-value is computed 4.1e-220 which is also significant.

REFERENCES:

1. Arabmazar, A. and Schmidt, P. 1981. Further evidence on the robustness of the Tobit estimator to heteroskedasticity. *Journal of Econometrics* 17, 253–8.
2. McDonald, J. F. and Moffitt, R. A. 1980. The Uses of Tobit Analysis. The Review of Economics and Statistics Vol 62(2): 318-321.
3. Olsen, R. 1978. A note on the uniqueness of the maximum likelihood estimator for the Tobit model. *Econometrica* 46, 1211–15.

4. Anderson, T. W., Kunitomo, Naoto, and Morimune, Kimio, "Comparing Single Equation Estimators in a Simultaneous Equation System," Technical Report No. 1, Econometric Workshop, Stanford University, January 1985.
5. Angrist, J. and J. Pischke (2009). Mostly harmless econometrics: An Empiricist's Companion. Princeton University Press.
6. Baltagi B. H., (2014), Econometric Analysis of Panel Data, 5th edition, Wiley.
7. Fuller, Wayne A., "Some Properties of a Modification of the Limited Information Estimator," Econometrica 45: 939-953.
8. Gallant, A. R. (1981), "On the Bias in Flexible Functional Forms and an Essentially Unbiased Form: The Fourier Flexible Form," Journal of Econometrics 15: 211–245.
9. Geweke, J. (1986), "Exact Inference in the Inequality Constrained Normal Linear Regression Model," Journal of Applied Econometrics 1: 127–142.
10. Greene, William. 2011. Econometrics Analysis, 7th ed. Prentice Hall, Upper Saddle River, NJ.
11. Hausman, J. and Wise, D. 1977. Social experimentation, truncated distributions and efficient estimation. *Econometrica* 45, 319–39.
12. Jeffrey M. Wooldridge, 2001. "Econometric Analysis of Cross Section and Panel Data," MIT Press Books, The MIT Press, edition 1, volume 1, number 0262232197, January.

