Spatial Tobit Regression Analysis: Case Study of Internet Use in Libya

Karima Hadia Abdulnabi
had71103@gmail.com

Jafara University/College of Education/Department of Mathematics

ABSTRACT

A special method is needed to analyze censored data that has a spatial correlation. If you use linear regression, it will result in an invalid parameter estimate, not being fulfilled assumption of normality and cloud the interpretation of the model. Spatial Tobit regression model was used to analyze data on internet usage in Libya. Parameter estimation using the method MCMC Gibbs sampler with Bayesian inferential approach. This study intends to establish a spatial Tobit regression model and to find a parameter estimation method from the spatial Tobit regression model. The data used as the response variable is the percentage of the population aged 5 years and over who accessed the internet during the last three months in the city of Tripoli, Libya in 2021. Censorship is given to districts/cities with a percentage of internet users greater than 16 percent, which is considered the minimum percentage of internet users to be achieved by the district / city.

Introduction

Using the classical linear regression model for analyzing censored data that has spatial correlation is an inappropriate decision. The term censored data is used to describe a group of data that has an unknown number of values at its upper or lower bound. Long (1997) explains that using a linear regression model on all censored data will produce parameter values that overestimate the slope and underestimate the intercept. Meanwhile, if you eliminate or cut observations whose values are unknown, it will produce parameter coefficients that underestimates the slope and overestimates the intercept. The truncated data causes a correlation between predictor variables and residuals, resulting in inconsistent estimates. Spatial correlation effects can appear in the formation of linear regression models that use cross-sectional data. This results in the failure to fulfill the assumptions of independent and identical errors that are normally distributed, resulting in invalid parameter estimates and obscuring the interpretation of the model (Marsh, Mittelhammer, & Huffaker, 2000). Spatial correlation can be observed from grouping certain values in data originating from adjacent areas, for example data on the level of internet usage in Libya.

The high level of internet usage in Libya is mainly found in big cities as centers of education and entertainment services, such as Tripoli. This spatial dependency phenomenon can be applied to the analysis of censored data that has a spatial correlation, where districts/cities with high category internet usage can be considered as data whose value is unknown. A special method is needed to

analyze internet usage, with the assumption that internet usage data in Libya is censored data that has a spatial correlation. Fischer and Getis (2010) said that censored data modeling involving areas should use spatial analysis, the most suitable method is spatial Tobit regression. In addition, Lee (2010) also stated that the spatial Tobit approach is more recommended for regional analysis involving censored data. Spatial Tobit regression analysis is used if the response variable in the spatial model involves data that is believed to have censored values (LeSage & Pace, 2009).

This study intends to form a spatial Tobit regression model and seek parameter estimation methods from the spatial Tobit regression model. The data used as a response variable is the percentage of the population aged 5 years and over who accessed the internet during the last three months on Libya in 2010. Censorship is given to districts/cities with a percentage of internet users greater than 16 percent, which is considered the minimum limit percentage of internet users that a country wants to reach.

Methodology

Tobit Regression Model

Suppose is a response variable with complete information and is sample data from, then the censored response variable can be defined as follows:

$$I(y) = \begin{cases} 0 & \text{if } y \leq y_L, \\ 1 & \text{if } y > y_L. \end{cases} \longrightarrow \quad (1)$$

Where , is a limiting constant, and is the number of observations (Tobin, 1958). If the value of yi is not known, then it contains a latent variable that cannot be observed in its entire range. depicted as a dark area on the curve. Panel A on shows the distribution. If the unknown value is cut, then some of the information that can explain the population will be lost where the distribution curve becomes more pointed (panel B). Panel C depicts censored data that is grouped on values so that it does not change information regarding population distribution.

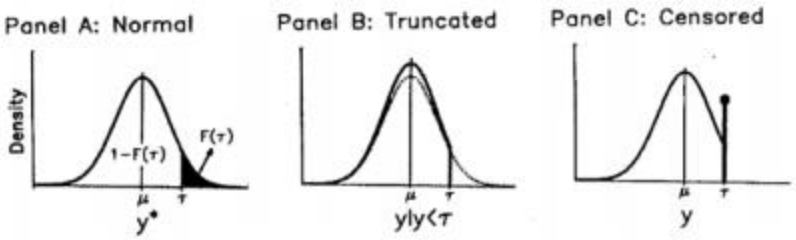

Fig. 1. The observations (Tobin, 1958)

Spatial Regression Models

In 1988, Anselin developed a general form of a spatial regression model (general spatial model) using cross section data as follows:

$$\mathbf{y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} (I - \lambda W)^{-1} \varepsilon} \longrightarrow \quad (2)$$

Where the formula is a response variable vector that has a spatial correlation, is a predictor variable matrix, and is a regression parameter vector. As for the spatial lag correlation coefficient of the response variable, is the spatial error correlation coefficient, and is a spatial weighing matrix with zero diagonal elements. This equation is also commonly referred to as a spatial autoregressive moving average (SARMA) regression model.

Markov Chain Monte Carlo (MCMC)

MCMC is a simulation method technique that generates a number of samples from known data distributions (Chib & Greenberg, 1996). The basic idea of the MCMC technique is that instead of calculating a probability density function, it is better to take a large random sample from it to find out the exact shape of the probability. With a large enough random sample size, the mean and standard deviation values can be calculated accurately (Casella & George 1992). LeSage (1999) explained that the MCMC Gibbs sampler algorithm will provide easy parameter estimation for a spatial Tobit regression model rather than having to solve a number of integral equations in the maximum likelihood method. The MCMC Gibbs sampler method aims to find the estimated value of using a conditional posterior distribution, where other values are assumed to be known. The posterior distribution of the parameters is determined by the principle of the Bayes theorem which is stated by:

$$p(\theta|\boldsymbol{y}) = \frac{L(\theta|\boldsymbol{y})\, p(\theta)}{m(\boldsymbol{y})} \propto L(\theta|\boldsymbol{y})\, p(\theta) \longrightarrow \quad (3)$$

One of the studies using the spatial Tobit regression model was conducted by Langyintuo and Mekuria (2008) who used the maximum likelihood method to form the Tobit SARMA model on farmer data in Mozambique. Kaliba (2002) developed the Tobit SARMA model using the Maximum Likelihood 4 application module from the GAUSS program package (developed by Aptech Systems, 1995) on rural data in Tanzania. Meanwhile LeSage and Pace (2009) used simulation data generated by Koop to form a spatial Tobit model using the Bayesian MCMC (Markov Chain Monte Carlo) approach with the Gibbs Sampling algorithm. As for research that uses the spatial Tobit model on information and communication technology (ICT) data, it has never been found.

ICT development of a country has a positive relationship with economic growth. That is, ICT development will have a chain effect on increasing economic growth (Kominfo, 2010). Rao and Pattnaik (2006) state that the growth of ICT has opened up opportunities for people to better utilize more modern socio-economic and cultural development facilities. ICT development provides a broad economic impact, both directly and indirectly, increasing welfare and building socio-economic facilities (ITU, 2010). Internet access is an indicator that best represents the level of ICT

development in a country, apart from economic growth in the telecommunications sector, cell phone ownership or computer control.

Using data from 154 countries, Howard and Mazaheri (2009) found that ICT usage gaps (cell phones, computers, and internet bandwidth) are influenced by; foreign investment, trade, population, urban population, literacy rate, consumption, wired telephone, and nine other variables that explain government regulation. Andonova and Serrano (2007) explain that the development of ICT and the growth in internet usage are more influenced by factors of government attention and regulations that apply in the region. Michailidis et al. revealed that internet users in rural Greece are influenced by income level, price of internet access, PC ownership, place of residence, as well as social demographic variables such as; gender, number of young people living in the same house, age, education level, and employment status (Michailidis, Partalidou, Nastis, Klavdianou, & Charatsari, 2011).

Based on the results of previous studies, the spatial Tobit regression model built in this study used the MCMC algorithm as a parameter estimation method. The level of internet usage in regencies/cities in Libya is used as a response variable with the following predictor variables; percentage of population living in urban areas, percentage of young people, percentage of population with high school graduates and above, average length of schooling, percentage of households with computers, percentage of households with cellular telephones, and percentage of villages with telephone signal cellular.

Analysis Method

Data Sources and Research Variables

The data source used in this research is processed data from the 2010 Susenas and 2008 Podes collected by the Central Bureau of Statistics (BPS). Spatial weighting matrices were prepared using the W queen contiguity method, namely districts/cities that border each other's areas will have a spatial correlation while those that are separated from each other will not have a correlation. The value if regions i and j border each other, becomes if they do not border each other. The digital map used is based on the results of updating the 2010 Population Census map.

The research object used as the response variable is the level of internet usage in 18 regencies/cities in Libya, namely the percentage of the population aged 5 years and over who have accessed the internet in the last three months. Censorship is given to districts/cities with a percentage of the population using the internet above 16 percent by considering the value to the value. The predictor variables used are as follows:

X1 : Percentage of population living in urban areas.
X2 : Percentage of young population (13-24 years).
X3 : Percentage of population who graduated from senior high school and above.
X4 : Average length of schooling.
X5 : Percentage of households that have computers.
X6 : Percentage of households that have cell phones.

X7 : Percentage of villages that have cell phone signal.

Spatial Tobit Regression Model

The spatial Tobit regression model is the application of the spatial regression model to censored data. So by combining equation (3) into (1), a general spatial Tobit regression model will be obtained as follows.

$$y_i = \begin{cases} y^* = (I - \rho W)^{-1} X\beta \\ \quad + (I - \rho W)^{-1}(I - \lambda W)^{-1}\varepsilon \\ \tau \end{cases} \longrightarrow \quad (4)$$

Complete Censored Data

Parameter estimation of the spatial Tobit regression model was carried out with the initial assumption that the response variable Y is data with complete information, uncensored, and has a spatial correlation. Even though according to equation (1), the data that has complete information is the response variable that follows a normal distribution. The value of when is an unknown observation or is considered a latent variable. So the value must be completed using a value.

Parameter Estimation

In accordance with the initial assumption that the response variable is data with complete information, not censored and has a spatial correlation, the relationship with the predictor variable is represented by a spatial regression model according to equation (4) above. LeSage (2000) and Lacombe (2008) formulated the conditional posterior distribution of each parameter as follows:

$$\begin{aligned} &\left(\frac{1}{\sigma^2}\boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{\varepsilon}\right) \sim \chi^2_{(n+4)} \\ &\left\{\frac{1}{v_{ii}}\left(\frac{\varepsilon_i^2}{\sigma^2} + r\right)\right\} \sim \chi^2_{(r+1)} \\ &p(\boldsymbol{\beta} \mid \sigma^2, \boldsymbol{V}, \rho, \lambda) \propto \\ &exp\left\{-\tfrac{1}{2\sigma^2}[\boldsymbol{B}(\boldsymbol{Ay} - \boldsymbol{X\beta})]^{\mathrm{T}}\, \boldsymbol{V}^{-1}[\boldsymbol{B}(\boldsymbol{Ay} - \boldsymbol{X\beta})]\right\} \\ &p(\rho, \lambda \mid \sigma^2, \boldsymbol{V}, \boldsymbol{\beta}) \propto |\boldsymbol{I} - \rho\boldsymbol{W}|\,|\boldsymbol{I} - \lambda\boldsymbol{W}| \\ &\qquad \times exp\left\{-\frac{1}{2\sigma^2}\boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{\varepsilon}\right\} \end{aligned} \longrightarrow \quad (5)$$

Parameter estimation of the MCMC Gibbs sampler method is carried out by generating random numbers that follow the conditional posterior distribution of each parameter for the desired number of iterations. The Metropolis within Gibbs algorithm is used in the non-standard form of the posterior distribution, namely for parameters or parameters (LeSage, 2000). To determine whether or not a predictor variable is feasible to be included in the model, the Wald test statistic is used with the following hypothesis:

$$\begin{aligned} &H_0: \beta_k = 0 \\ &H_1: \beta_k \neq 0,\ k = 1,2,\ldots,p \end{aligned} \longrightarrow \quad (6)$$

Results and Discussion

Internet use in Libya

The level of internet use is seen from the percentage of the population aged five years and over who have accessed the internet in the last three months. The average internet usage per district/city in Libya is 12.02%. Regions with internet usage values around the average are Bani Walid, Surt, Chadamis, Zuwarah, and Al-Aziziyah, with 2.23% and 2.54% respectively.

Fig. 2. Internet usage in Libya

The thematic map of internet use in Figure 2 shows the high use of the internet in the cities as student, industrial and business centers was followed by other areas around it. Areas directly in contact with these cities have a slightly lower percentage of internet users, while the next area that intersects indirectly has an even lower percentage.

For spatial Tobit regression modeling, the percentage value of internet users from 34 districts/cities is considered unknown. They are areas with a level of internet usage greater than 16 percent, which is the minimum limit for the percentage of internet users that a district/city wants to achieve. The values of the unknown level of internet use are considered equal to 16 percent, so that the variable percentage of internet users is obtained as censored data. This is in accordance with the concept in equation (1) above.

Descriptive data on the level of internet usage as a censored response variable can be seen in Table 1 below. The maximum value of the percentage of internet users per district/city is 16 percent, with an average and standard deviation of 9.97 percent and 4.63, respectively. The variable that has the greatest variation in value is the percentage of urban population with a standard deviation of 30.84 and the length of the data range from 9.27 percent to 100 percent. The average length of school variable has the smallest variation with a standard deviation of 1.52. This is because the units of these variables are in years, while other variables are in percentage units.

Table 1. Descriptive Research Variables

Research variable	Deskription	Min	Max	Average	Standard Deviation
(1)	(2)	(3)	(4)	(5)	(6)
Y	Percentage of internet users	2,23	16,00	9,968	4,626
X_1	Percentage of urban population	9,27	100,00	57,991	30,837
X_2	Percentage of young population	12,44	25,00	17,395	2,567
X_3	Percentage of population graduating from high school and above	5,30	50,26	21,349	11,105
X_4	Average length of school	4,21	11,55	8,052	1,518
X_5	Percentage of households that have a computer	2,20	40,22	11,433	9,127
X_6	Percentage of households owning a cell phone	44,87	94,89	71,923	12,050
X_7	Percentage of villages with telephone signal	50,00	100,00	88,892	10,026

Multiple linear regression modeling using the ordinary least squared (OLS) method was carried out to explain the relationship between predictor variables on the level of internet use in Libya. At a 95 percent degree of confidence, the parameter test results yield only two of the seven predictor variables that affect the response variable. The variance inflation factor (VIF) value is very high for the variable and indicates the existence of multicollinearity conditions between predictor variables. Although it produces a fit model with a very high coefficient of determination, the multiple linear regression model obtained is not appropriate to use as a basis for analysis. This is due to non-fulfillment of the non-multicollinearity assumption and the amount of wasted information from predictor variables.

Table 2. Multiple Linear Regression Model and Variance Inflation Factor (VIF) Values

Parameter	Koefisien	Parameter Test		VIF Statistics
		Test Statistics t	*p-Value*	
(1)	(2)	(3)	(4)	(5)
$\hat{\beta}_0$	-13,729	-4,158	0,000	–
$\hat{\beta}_1$	0,022	1,960	0,053	5,705
$\hat{\beta}_2$	-0,022	-0,298	0,766	1,701
$\hat{\beta}_3$	0,107	1,585	0,116	27,368
$\hat{\beta}_4$	0,797	1,936	0,055	18,952

	-0,026	-0,531	0,597	9,827
$\hat{\beta}_5$	0,083	2,671	0,009	6,872
$\hat{\beta}_6$	0,094	4,582	0,000	2,072
$\hat{\beta}_7$				
Analysis of Variance– (uji)		132,47	0,000	–
Koefisien Determinasi (R^2)	0,894	–	–	–

According to Gujarati (2004), multicollinearity conditions can be handled by selecting variables, either by adding new predictor variables or by reducing existing ones. Another way that can be done is to use another model that is more appropriate to explain the relationship between predictor variables and responses. The spatial Tobit regression model is more appropriate to use to explain the factors that influence the high and low diversity of internet use among districts/cities in Libya which are influenced spatially by the surrounding areas.

The predictor variables that were excluded from the model were (percentage of young people) and (percentage of households with computers). In the final stage of backward elimination, five predictor variables were obtained that significantly influenced the diversity of internet use in Libya. Based on equation (21) and the results of parameter estimation presented in table 3 above, the Tobit spatial lag regression model that is formed is:

$$\hat{y}_i = \begin{cases} -13,727 - 0,17 \sum_{j=1, j\neq i}^{n} w_{ij} y_j + 0,012 x_{1i} \\ +0,144 x_{3i} + 0,755 x_{4i} + 0,093 x_{6i} + 0,083 x_{7i} \; ; \; y_i < 16 \\ 16 \qquad\qquad\qquad\qquad\qquad\qquad ; \; y_i \geq 16 \end{cases} \longrightarrow \quad (7)$$

This model is used to explain the factors and surrounding areas that affect the level of internet usage in a district/city, when the value is less than 16 percent. As for regencies/cities with high levels of internet usage, they are considered as benchmarks for ICT development to be achieved.

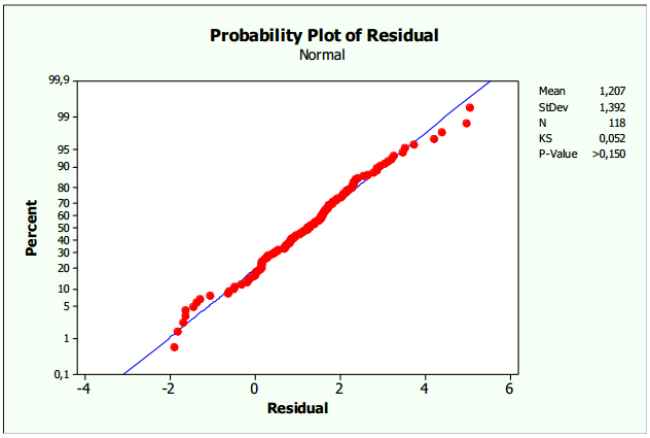

Figure 3. Graph of Normal Probability Plot of Residual Spatial Lag Tobit Regression Model

The normal probability plot graph of the residuals is used to test the normality assumption of the model error. The residual plot appears to be around the normal probability line, so it can be concluded that the normality assumption of the model error is significantly fulfilled. The next assumption regarding the problem of homogeneous error variance is considered to have been met, considering the results of the previous Breusch Pagan test which showed that the heteroscedasticity condition was not met and the MCMC simulation process was based on the homoscedasticity condition. The assumption that there is no autocorrelation in errors has also been fulfilled, because the model formed is not a Tobit spatial error regression model.

Model Interpretation

The above equation explains that for districts/cities with a percentage of internet users that is less than 16 percent, the level of internet usage in that area is influenced by other areas in the vicinity and the following variables:

- Percentage of population living in urban areas (X1). If other variables are held constant, then everyone percent increase in the percentage of people living in urban areas will result in an increase in the percentage of internet users by 0.012 percent.
- Percentage of population graduating from high school and above (X3). If other variables are held constant, then everyone percent increase in the percentage of high school graduates and above will result in an increase in the percentage of internet users by 0.144 percent.
- Average length of schooling (X4). If other variables are held constant, then each increase in the average length of schooling for one year will result in an increase in internet usage for a district/city by 0.755 percent.
- Percentage of households that have a cell phone (X6). If other variables are held constant, then everyone percent increase in the percentage of households owning a mobile phone will result in an increase in the percentage of internet users by 0.093 percent.

- Percentage of villages that have cell phone signal (X7). If other variables are held constant, then everyone percent increase in the percentage of villages that receive cell phone signals will cause an increase in internet use by 0.083 percent.

The coefficient of determination shows that 83.94 percent of the variation in internet use in Libya is explained by the five predictor variables in the model, the rest by other variables. The effect of the spatial lag from other areas that intersect areas can be seen in the Tobit spatial lag regression model for each district/city, namely the y1 model when. As for districts/cities with high internet usage, the percentage of population aged 5 years and over who has accessed the internet in the last three months is considered to be 16 percent or y1 = 16 when yi ≥ 16.

The percentage of population living in urban areas indicates the level of progress and completeness of public facilities in the area. The percentage of the population who have graduated from high school and above and the average length of schooling reflect the quality of human resources in the area. Thus, increasing the percentage of internet users can be done through efforts to improve the quality of human resources from the educational aspect. In addition, the construction of complete public facilities in rural areas can also increase the level of internet usage in districts/cities.

The variable characteristics of equipment and networks in the regions indicate the importance of the development of cellular telephone technology for the growth of the internet. The various ease of internet access provided by mobile devices and the breadth of the cellular telephone network have significantly boosted the level of internet usage. On the other hand, the use of the internet is not actually accessed via computers or is dominated by a young population. The internet can be accessed by anyone and through any media, especially cell phones.

Apart from being influenced by the five variables above, the level of district/city internet usage in Libya is also influenced by other regions that intersect with the region. For example, the level of internet usage in Libya can be explained through the Tobit spatial lag regression model below:

$$\hat{y}_{3101} = -0{,}085(y_{3175} + y_{3603}) + \boldsymbol{x}_{3101}{}^{T}\boldsymbol{\beta} \longrightarrow \quad (8)$$

Where is the predictor variable vector from the parameter vector. The level of internet usage in Libya is also influenced by internet usage in Tripoli. If other variables are held constant, internet usage in the Libya is -0.085 times that of the combined internet usage in Libya. In detail, each of the 84 Tobit regression models has spatial lag when the percentage value of internet users is less than 16%.

Conclusions And Recommendations

The spatial Tobit regression model is a spatial regression model that is applied to censored data, with the general model form of spatial Tobit

regression being:

$$y_i = \begin{cases} \rho w_i^T y + \lambda w_i^T y + \left(\rho W^T \lambda W\right)_i^T y \\ +x_i^T \beta - \lambda\left(w_i^T X\beta\right) + \varepsilon_i \\ \tau \end{cases} \longrightarrow \quad (9)$$

The parameter estimation method that can be used is Markov Chain Monte Carlo (MCMC) which is equipped with the Gibbs sampler and Metropolis within Gibbs algorithms. This method puts forward computational simulation techniques to generate a large number of random variables using the Bayesian inference approach. Using the use of the internet in Libya as a case study, it is known that the Tobit Spatial lag regression model produces richer information than the multiple linear regression model. Factors that affect district/city internet use in Libya are the percentage of the population living in urban areas, the percentage of the population who have graduated from senior high school and above, the average length of schooling, the percentage of households that have cell phones, and the percentage of villages that have access to internet. Cell phone signal. Apart from being influenced by these five variables, the level of district/city internet use in Libya is also influenced by other regions that intersect with the region.

Based on the research results that have been obtained, further development can be carried out using the highest posterior density (HPD) and Bayes factor as a method of testing parameters and models. This study still uses a queen contiguity weighing matrix, so that in future studies it can be developed using other weighing matrices such as distance. Furthermore, the MCMC Gibbs sampler method for spatial Tobit regression modelling can be used for other data and cases that are more applicable.

References

Andonova, V., & Serrano, L. D. 2007. *Political Institutions and the Development of Telecommunications.* Bonn: IZA Discussion Paper. Anselin, L. 1988. *Spatial Econometrics: Methods and Models.* Dordrecht: Kluwer Academic Publishers.

Anselin, L. 1999. *Spatial Econometrics.* Dallas: University of Texas.

Breusch, T., & Pagan, A. 1979. A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica, Vol. 47, No. 5,* 1287-1294.

Casella, G. dan Berger, R. 2002. *Statistical Inference.* Duxbury, Thomson Learning.

Casella, G. dan George, E. I. 1992. Explaining the Gibbs Sampler. *The American Statistician, Vol. 46, No. 3,* 167-335.

Chib, S. dan Greenberg, E. 1996. Markov Chain Monte Carlo Simulation Methods in Econometrics. *Econometrics Theory, Vol. 12,* 409-431.

DeMaris, A. 2004. *Regression with Social Data: Modelling Continuous and Limited Response Variable.* New Jersey: John Wiley and Sons, Inc.

Draper, N. R. dan Smith, H. 1998. *Applied Regression Analysis.* New York: John Willey and Sons, Inc.

Fischer, M. M. dan Getis, A. 2010. *Handbook of Applied Spatial Analysis: Software Tools, Methods, and Application.* New York: Springer.

Greene, W. H. 2008. *Econometric Analysis, Sixth Edition.* New York: Pearson - Prentice Hall.

Hastings, W. 1970. Monte Carlo Sampling Methods using Markov Chains and Their Applications. *Biometrika, Vol. 57, No. 1*, 97-109.

Howard, P. N. dan Mazaheri, N. 2009. Telecommunications Reform, Internet Use, and Mobile Phone Adoption in Developing World. *World Development, Vol. 37, No. 7*, 1159-1169.

Kaliba, A. R. 2002. Participatory Evaluation of Community Based Water and Sanitation Programes: The Case of Central Tanzania. *Dissertation.* Mahattan: Kansas State University.

Langyintuo, A. S. dan Mekuria, M. 2008. Assessing the Influence of Neighborhood Effects on the Adoption of Improved Agricultural Technologies in Developing Agriculture. *AfJARE, Vol. 2, No. 2*, 151-169.

Lee, M. J. 2010. *Micro-Econometrics: Methods of Moments and Limited Dependent Variables, Second Edition.* New York: Springer.

LeSage, J. P. 1999. *The Theory and Practice of Spatial Econometrics.* Ohio: University of Toledo.

LeSage, J. P. 2000. Bayesian Estimation of Limited Dependent Variable Spatial Autoregressive Models. *Geographical Analysis, Vol. 32, No. 1*, 19-35.

Long, J. S. 1997. *Regression Models for Categorical and Limited Dependent Variables.* California: Sage Publications, Inc.

Marsh, T. L., Mittelhammer, R. C., & Huffaker, R. G. 2000. Probit with Spatial Correlation by Field Plot: Potato Leafroll Virus Net Necrosis in Potatoes. *Journal of Agricultural, Biological, and Environmental Statistics*, Volume 5, Number 1, Pages 22-36.

Michailidis, A., Partalidou, M., Nastis, S. A., Klavdianou, A. P.dan Charatsari, C. 2011. Who Goes Online? Evidence of Internet Use Patterns from Rural Greece. *Telecommunications Policy, Vol. 35*, 333-343.

Rao, J. G. dan Pattnaik, S. 2006. Technology for Rural Development Role of Telecommunication Media in India. *Indian Media Studies Journal, Vol. 1, No. 1*, 85-92.

Tobin, J. 1958. Estimation of Relationships for Limited Dependent Variables. *Econometrica, Vol. 26, No. 1*, 24-36.