IJRP.ORG
International Journal of Research Publications
ISSN: 2708-3578 (Online)

An Estimator For Simultaneous Equation Model Using Two Stage Ridge Method (Case Study On Farmer Exchange Rate Data In Indonesia)

Sarina[a*], Nurtiti Sunusi[a**], Erna Tri Herdiani[a***]

[*] *Corresponding Author: sarinar328@gmail.com*
[**] *Corresponding Author nurtitisunusi@unhas.ac.id*
[***] *ernatriherdiani@unhas.ac.id*
[a]*Departement of Statistics, Hasanuddin University, Makassar, 90245, Indonesia*

Abstract

The concept of simultaneity is undoubtedly the most influential idea in econometrics, such as the relationship in farmer exchange rate. A simultaneous equations model is model with two or more equations is defined as one in which a variable explained in one equation appears as an explanatory in another. In the simultaneous equation model, the variables used are known as endogenous variables and exogenous variables. As a result,the model's endogenous variables are determined at the same time. Moreover, in these simultaneous equations, there is a correlation between the error terms of the structural equations of the model, the two stage least square method was selected to estimate. Under the issues of multicollinearity two stage least squares estimation in a simultaneous equations model has several desirable properties. Futhermore, we use a two Stages least square estimator for the simultaneous equations model, which suffers from autocorrelation issues and then we combined with ridge regression estimator which suffers from multicollinearity issues. After adjusting this with the ordinary ridge regression estimator, we use a mixed method to apply the two stages least squares procedure. From the this study, it was found that the two stage ridge is better than two stage least square. This is influenced by the simultaneous relationship and multicollinearity in each equation.

Keywords: Simultaneous Equation Models; Multicollinearity; Two Stage Least Square; Ridge Regression

1. Introduction

The dependent variable in two or more equations is also the independent variable in several other equations in a simultaneous equation (Koutsoyiannis, 1978). Thus, a variable in the simultaneous equation has two roles, namely as endogenous variable and as exogenous variable (Supranto, 1984). In simultaneous equations, a set of endogenous variables to a set of exogenous variables with error variables and the terms endogenous with error variables. In this method, endogenous variables can become exogenous variables that endogenous variables have a possibility to relate exogenous variables. Endogenous variable are variables whose value are determined in the system of equations, while exogenous variables are determined outside the model (Syafaat, 1996).

In general, an equation is thought to represent a relationship describing a phenomenon in any regression modelling. Many situations involve a set relationships that explain how certain variables behave. As a result, parameter estimation in this situation has characteristics that are not present when a model involves only a single equation. When a relationship is a component af a system, some explanatory variables are stochastic and correlated with the error term. In investigating the realationship between explanatory and response variables in a single equation. Ordinary Least Squares (OLS) is the best linear unbiased estimator (BLUE) in investigating the relationship between explanatory and response variables in a single equation. The OLS is only applicable when all regression assumptions are satisfied and some of them are; errors in the model are distributed withnormal distribution with zero mean and a constant and nohigh correlation problem among the explanatory (independent variables) (Shariff and Duzan, 2018). However, the endogenous variable in the simultaneous equation are correlated with error (disturbance) so OLS estimator will produce a biased and inconsistent (Eledum and Alkhalifa, 2012).

Several alternative for estimating simultaneous equation which is associated with the autocorrelation of errors which occurs when the value of the error term in any particular period is correlated with its own preceding value or values are the reduce form equations, two stage least square estiamation (2SLS), indirect least square, and three stage least square. The 2SLS method was introduced by (Theil, 1953), (Basmann, 1957). The 2SLS better than ILS, because it gets one estimator for one parameter and returns a standard error for each estimator (Gujarati, 2010), (Lopez-Espin, Vidal., and Giménez, 2012)

In the simultaneous equations, the problem of multicollinearity may still exist in the individual equations. If the simultaneous equation solution to this problem is adopted there may be an intolerable rise in the size of the model with the consequent depletion of the number of exogenous variables (Miftahul, Alfian, & Dian, 2011). Multicollinearity is defined as conditions on which some or all explanatory variables have large influence on others explanatory variables. This critical issue might happen if the analysis contains large sets of data with several numbers of explanatory variables and this will affect to the existence of multicollinearity problem. In the presence of multicollinearity, the regression assumptions are invalid and as such, the OLS cannot be preceded in the next stage of estimation. Otherwise, the results of parameter estimates and inference under OLS procedure will be insignificant and unreliable. Due to such problem, there are quite number methods of estimations to overcome multicollinearity problem in regression analysis. Hoerl and Kennard (1970) are the first to introduce ridge regression method by adding small positive quantities (denoted by letter k in many studies to the diagonal of the matrix $X'X$ where $X'X$ is matrix of explanatory variables) and it is shown can minimize the biased estimates and mean squared error (MSE) of the model (Duzan, H., Shariff, 2015).

In this paper, we purpose a method for estimating the parameters for mixed problem, namely auto-correlated errors and multicollinearity. We chose to combined the two stage least square estimator and ridge regression estimator. This study proposes another technique with k can be formed as a linear combination of coefficients of determination of explanatory variables (Mansson, Shukur, Kibria, 2010). The performance of the proposed method is investigated and comparison is made to 2SLS and some existing methods by using Variance Inflation Factor (VIF) and MSE criterion. In this article, what is presented is a description of the estimation two stage ridge regression.

2. Material

This section will explain the research data, sampling techniques, and research variables.

2. 1 Two Stage Least Square for Simultaneous Equation

Each simultaneous equation is composed of two variables, namely exogenous and endogenous variables.

Endogenous variables are dependent variables whose value are determined in the simultaneous equations. Exogenous variables are variables whose values have been determined outside the model. The equations in the model are called structural equations while the parameters are called structural parameters. Structural parameters reflect the direct effect of each exogenous variable on endogenous variables. A simultaneous model is said to be complete if the number of equations in the system likely the number of endogenous variables. An estimator called two-stage least squares would involve running OLS two times. Assume we want to estimate the coefficients of the linear model

$$Y_i = \beta_0 = \beta_1 X_{1i} + \cdots + \beta_p X_{pi} + \varepsilon_i \qquad (1)$$

The two stage least square estimator of β is the following procedure:

1. First Stage: Z affects X

Regress each X_j on Z and save the predict values, $\widehat{X}_j$. If X_j is included in Z, we will have $\widehat{X}_j = X_j$. Each endogenous variable is regressed against all exogenous variable of a system so that the equation of the reduced form is obtained. Suppose there is multiple regression equation:

$$Y_1 = \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \cdots + \beta_p X_{1p} + \varepsilon_1$$
$$Y_2 = \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \cdots + \beta_p X_{2p} + \varepsilon_2$$
$$\vdots$$
$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} + \varepsilon_i$$

But some of the variable X_{ij} are correlated with the error term.

OLS estimation of this equation will be biased and inconsistent. Where equation (1) is written in matrix form as follows

$$Y_i = X\beta + \varepsilon_i \qquad (2)$$

From equation (2), got the following error as:

$$\varepsilon_i = Y_i - X\beta \qquad (3)$$

To minimize equation (3), it is obtained by finding the derrivative of the error (ε) with respect to β and then equating each dervative to zero

$$\frac{\partial \varepsilon}{\partial \beta_0} = -2\sum \left(Y_i - \beta_0 - \beta_1 X_{1p} - \beta_2 X_{i2} - \cdots - \beta_p X_{ip}\right) = 0$$
$$\frac{\partial \varepsilon}{\partial \beta_1} = -2\sum \left(Y_i - \beta_0 - \beta_1 X_{1p} - \beta_2 X_{i2} - \cdots - \beta_p X_{ip}\right) = 0$$
$$\frac{\partial \varepsilon}{\partial \beta_2} = -2\sum \left(Y_i - \beta_0 - \beta_1 X_{1p} - \beta_2 X_{i2} - \cdots - \beta_p X_{ip}\right) = 0$$
$$\frac{\partial \varepsilon}{\partial \beta_p} = -2\sum \left(Y_i - \beta_0 - \beta_1 X_{1p} - \beta_2 X_{i2} - \cdots - \beta_p X_{ip}\right) = 0$$

Let $\beta_1, \beta_2, \cdots, \beta_p$ dinyatakan dengan $b_1, b_2, \cdots, b_k$, maka

$$\sum Y_i \quad = nb_0 + b_1 \sum X_{i1} + b_2 \sum X_{i2} + \cdots + +b_p \sum X_{ip}$$
$$\sum Y_i X_{i1} = b_0 X_{i1} + b_1 \sum X_{i1}^2 + b_2 \sum X_{i2} X_{i1} + \cdots + +b_p \sum X_{ip} X_{i1}$$
$$\sum Y_i X_{i2} = b_0 X_{i2} + b_1 \sum X_{i1} X_{i2} + b_2 \sum X_{i2}^2 + \cdots + +b_p \sum X_{ip} X_{i2}$$
$$\sum Y_i X_{ip} = b_0 X_{ip} + b_1 \sum X_{i1} X_{ip} + b_2 \sum \mathrm{X_{i2} X_{ip}} + \cdots + \mathrm{b_p} \sum \mathrm{X_{ip}^2} \qquad (4)$$

Equation (4) can write $(X'X)b = X'Y$ so that the estimator b is:

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{5}$$

then look for the X projection to the Z projection which is the instrument variable matrix.

$$\begin{aligned}\hat{X} &= Z\hat{\beta} \\ &= (Z'Z)^{-1}Z'X \\ &= P_Z X\end{aligned} \tag{6}$$

2. The second Step

Estimate β via the OLS estimate of the regression model Eq. (1)

$$Y = \hat{X}\beta + \varepsilon$$

$$\varepsilon = Y - \hat{X}\beta$$

$$\begin{aligned}\varepsilon'\varepsilon &= \left(Y - \hat{X}\beta\right)'\left(Y - \hat{X}\beta\right) \\ &= \left(Y' - \hat{X}'\beta'\right)\left(Y - \hat{X}\beta\right) \\ &= \left(Y'Y - Y'\hat{X}\,\beta\right) - \left(Y\beta'\hat{X}' + \beta'\hat{X}'\hat{X}\beta\right)\end{aligned} \tag{7}$$

$$\frac{\partial\varepsilon'\varepsilon}{\partial\beta} = 0$$

$$\frac{\partial\varepsilon'\varepsilon}{\partial\beta} = -2\hat{X}'Y + 2\hat{X}'\hat{X}\,\beta$$

$$0 = -2\hat{X}'Y + 2\hat{X}'\hat{X}\,\beta$$

$$2\hat{X}'\hat{X}\,\beta = 2\hat{X}'Y$$

Where $E(\varepsilon^*) = 0$ and $Cov(\varepsilon^*) = \sigma^2 I_n$, Therefore, the OLS estimator for the model Eq. (7) is:

$$b = \left(X^{*'}X^*\right)^{-1}X^{*\prime}Y^* \tag{8}$$

Where

$$Y^* := P_Z Y = \begin{pmatrix} \sqrt{1-P_Z{}^2} & 0 & 0 & \cdots & 0 \\ -P_Z & 1 & 0 & \cdots & 0 \\ 0 & -P_Z & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -P_Z & 1 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ \vdots \\ Y_n \end{pmatrix}$$

$$X^* := P_Z X = \begin{pmatrix} \sqrt{1-P_Z{}^2} & 0 & 0 & \cdots & 0 \\ -P_Z & 1 & 0 & \cdots & 0 \\ 0 & -P_Z & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -P_Z & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ \vdots \\ X_n \end{pmatrix}$$

Note that $X^{*\prime}X = X'P_Z{}'P_Z X = X'\Omega X$ and $X^{*\prime}Y = X'P_Z{}'P_Z Y = X'\Omega Y$, where:

$$\Omega = P_Z{}'P_Z := \begin{pmatrix} 1 & -P_Z & 0 & \cdots & \cdots & 0 \\ -P_Z & 1+P_Z{}^2 & -P_Z & \ddots & \cdots & 0 \\ 0 & -P_Z & 1+P_Z{}^2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1+P_Z{}^2 & -P_Z \\ 0 & \cdots & \cdots & 0 & -P_Z & 1 \end{pmatrix}$$

Thus, 2SLS estimator is given as

$$b_{2sls} = (X'\Omega X)^{-1}X'\Omega Y \quad (9)$$

2.2 Ridge Regression Model

The 2SLS estimate in equation (9) is invalid due to large deviation in b. The $\hat{b}$ is said to be unbiased but inconsistent. To overcome this problem, the positive value of k is added to the diagonal elements in $X'\Omega X$ matrix in Eq. (9) to minimize the impact of high correlation in explanatory variables [14]. Hence, the ridge regression model for simultaneous equation is:

$$\hat{b}_R = (X'X + kI)^{-1}X'Y \quad (10)$$

According to [11], the value of the ridge estimator k in Eq. (10) will provide a smaller value of MSE compared to OLS estimator.

2.3 Two Stage Ridge Regression

In this study, the performance of existing ridge estimator are compared with 2SLS and the proposed simultan equation model with multicollinearity via two stage ridge regression estimator. According to equation (9) with $X^{*'}\Omega X^*$ is a symmetric matrix, so there is on orthogonal T, such that: $T'(X^{*'}\Omega X^*)T = \Lambda$

$$T'X^{*'}\Omega X^*T = \Lambda$$
$$(X^*T)'\Omega X^*T = \Lambda$$
$$M'\Omega M = \Lambda \quad (11)$$

Where Λ is diagonal matrix of size $q \times q$ with the elements of the main diagonal being the eigenvalues $(\lambda_1, \lambda_2, \ldots, \lambda_q)$ of the orthogonal matrix $X^{*'}\Omega X^*$ whose elements are the eigenvector values of $X^{*'}\Omega X^*$. Based on equation (11), then equation (2) can be written as follows:

$$\begin{aligned} Y^* &= X^*\beta + \varepsilon^* \\ &= X^*I\beta + \varepsilon^* \\ &= X^*TT'\beta + \varepsilon^* \\ &= (X^*T)T'\beta + \varepsilon^* \\ &= M\alpha + \varepsilon^* \end{aligned} \quad (12)$$

with $M = (X^*T)$ dan $\alpha = T'\beta$

from (12) we get:

$$\varepsilon^* = Y^* - M\alpha$$

$$\begin{aligned} H &= \varepsilon^{*'\varepsilon^*} \\ &= (Y^* - M\alpha)'(Y^* - M\alpha) \\ &= (Y^{*'} - M'\alpha')(Y^* - M\alpha) \\ &= (Y^{*'}Y^* - Y^{*'}M'\alpha' - \alpha'M'Y^* + \alpha'M'M\alpha) \\ &= (Y^{*'}Y^* - 2\alpha'M'Y^* + \alpha'M'M\alpha) \end{aligned}$$

$$\frac{\partial H}{\partial \alpha} = 0$$

$$-2M'Y^* + 2M'M\alpha = 0$$
$$M'M\alpha = M'Y^*$$
$$\hat{\alpha} = (M'M)^{-1}M'Y^*$$
$$= \Gamma^{-1}M'Y^* \quad (13)$$

from equation (13) can be formed into

$$\hat{\alpha} = \left(\mathrm{T'X^{*'}}\Omega XT\right)^{-1}(X^*T)'Y^*$$
$$\hat{\alpha} = (\mathrm{T'X^{*'}}\Omega XT)^{-1}T'X^{*'}X^*\hat{\beta}$$
$$\hat{\alpha} = (\mathrm{T'X^{*'}}\Omega XT)^{-1}T'X^{*'}X^*I\hat{\beta}$$

$\hat{\alpha} = (\mathrm{T'X^{*\prime}}\Omega XT)^{-1}T'X^{*\prime}X^{*}TT'\hat{\beta}$
$\hat{\alpha} = (\mathrm{T'X^{*\prime}}\Omega XT)^{-1}(T'X^{*\prime}X^{*}T)T'\hat{\beta}$
$\hat{\alpha} = T'\hat{\beta}$
$\hat{\beta} = T\hat{\alpha}$ *(14)*

by using the lagrange multplier, where $\hat{\alpha}$ is the value that minimize the objective function with the constraining condition:

$\alpha'\alpha - c^2 = 0$

obtained

$$H = (Y^* - M\alpha)'(Y^* - M\alpha) + k(\alpha'\alpha - c^2)$$
$$= Y^{*\prime}Y^* - Y^{*\prime}M\alpha - \alpha'M'Y^* + \alpha'M'M\alpha + k(\alpha'\alpha - c^2)$$
$$= Y^{*\prime}Y^* - 2\alpha'M'Y^* + \alpha'M'M\alpha + k\alpha'\alpha - kc^2) \quad (15)$$

qualified:

$\left.\frac{\partial H}{\partial \alpha}\right|_{\alpha=\hat{\alpha}} = 0$

So that

$$0 = 0 - 2M'Y^* + 2M'M\hat{\alpha} + 2k\hat{\alpha}$$
$$0 = -M'Y^* + M'M\hat{\alpha} + k\hat{\alpha}$$
$$M'Y^* = M'M\hat{\alpha} + k\hat{\alpha}$$

where k is a constant, therefore

$$M'M\alpha + k\alpha = -M'Y^*$$
$$(M'M\alpha + k)\alpha = -M'Y^*$$
$$\alpha = (M'M + k)^{-1}M'Y^*$$
$$= (((X^*T)'(X^*T) + K)^{-1}(X^*T)'Y^*)$$
$$= ((P_Z XT)'(P_Z XT) + K)^{-1}(\rho XT)'\rho Y)$$
$$= ((T'X'P_Z{}'P_Z XT + K)^{-1}T'X'P_Z'P_Z Y)$$
$$= ((T'X'\Omega XT + kI)^{-1}T'X'\Omega Y)$$

Therefore, the two stages ridge regression estimator is

$$\beta_{TR} = ((T'X'\Omega XT + kI)^{-1}T'X'\Omega Y) \quad (16)$$

Where k is a constant and Ωas defined in Eq. (9). Estimator two stage ridge is a bias estimator and its expectation is given as

$$E(\beta_{TR}) = \beta - kI((T'X'\Omega XT + kI)^{-1}\beta \quad (17)$$

3. Method

3.1 Data Sources and Research Variables

The data used in the application of the paramater estimation results was data on price index received by farmers and price index paid farmers. This data sourced from website www.bps.go.id in January 2011-December 2016. Completely the research variables can be seen on table 1.

Table 1. List of Variable

Variable	Variable Name
Endogenous	Y_1 = price index received by farmers (%) Y_2 = price index paid farmers (%)
Exogenous	X_1 = Producer Price Index (%) X_2 = Household Consumption Index (%) X_3 =Production Cost Index and Addition of Capital Goods (%)

3.2 Step of Analysis

The analysis steps are as follows:
1. Test classic assumptions and detect multicollinearity by looking at the VIF value of the independent variable.
2. Forming a reduction equation based on a structural model
3. Perform the Simultaneous Test to see that every equation have a simultaneous relationship between equation
4. Identify the model using the order condition method
5. Estimating the parameters of the simultaneous equation model using the 2SLS method.
6. Perform data transformation using centering and scaling for response variables and predictor variables
7. Perform the orthogonalization process on the independent variables by multiplying the independent variable X by the orthogonal eigenvectors.
8. Determine the initial estimator of the 2SLS method for equations that experience multicollinearity
9. Estimating the parameters of the two stage ridge
10. Calculating the mean square error (MSE) and variance of the two stage ridge parameter estimators for the model suitability test

4. Result

4.1 Description Data

For farmer exchange rate data which consists of two endogenous variables and three exogenous variables, the model specifications are as follows:

$$Y_1 = \gamma_{10} + \gamma_{12}Y_2 + \beta_{11}X_1 + \varepsilon_1 \qquad (18)$$

$$Y_2 = \gamma_{20} + \gamma_{21}Y_1 + \beta_{12}X_2 + \beta_{23}X_3 + \varepsilon_2 \qquad (19)$$

Based on minitab output, the VIF value for equation (18) for each exogenous variable is shown in Table 2 dan the VIF value for equation (19) for each exogenous variable is shown in Table 3.

Table 2. The VIF Value Exogeneous Variable for It

Variable	The VIF Value	Description
Y_2	1.364	Non Multicollinearity
X_1	1.364	Non Multicollinearity

Table 3. The VIF Value Exogeneous Variable for Ib

Variable	The VIF Value	Description
Y_1	224.278	Multicollinearity
X_2	55.596	Multicollinearity
X_3	115.801	Multicollinearity

Tabel 2 shows that all exogenous variables have a value $VIF < 10$. Thus, it can be concluded that there is no multicollinearity problem in It data. While from table 3, it can be seen that all exogenous variables have a VIF value > 10 like $Y_1 = 224.278, X_2 = 55.596$ and $X_3 = 115.801$. This shows that there is multicollinearity problem in the Ib data.

Tabel 4. Simultaneous Test

Model	Error	F	p_{value}	Description
Eq.(18)	ε_2	6947.6	0.000	Simultan
Eq.(19)	ε_1	399079.2	0.000	Simultan

The significanse of the error variable in each equation is shown in table 4.In equation (18) and (19) with significant error α = 0.05 indicates that there is a simultaneous effect between equations in the model. Simultaneous test results show that both of equation contain simultaneous effects. Therefore, simultaneous parameter estimation can be using 2SLS method.

4.2 The Application of The Parameter Estimation Two Stage Least Square

The First Stage

Apply OLS estimator to equation (18) and equation (19). Based on the calculation of the matlab program, the coefficients for each variable rduced form:

$$\begin{bmatrix}\pi_{10}\\\pi_{11}\\\pi_{12}\\\pi_{13}\end{bmatrix}=\begin{bmatrix}-30.1710\\-0.0004\\0.5517\\0.7868\end{bmatrix}\text{ dan }\begin{bmatrix}\pi_{20}\\\pi_{21}\\\pi_{22}\\\pi_{23}\end{bmatrix}=\begin{bmatrix}-0.0062\\-0.0003\\0.7442\\0.2738\end{bmatrix}$$

Then

$$\hat{y}_1=-30.2-0.0004\,X_1+0.5521\,X_2+0.7868\,X_3 \quad (20)$$
$$\hat{y}_2=-0.0062-0.0003\,X_1+0.7442\,X_2+0.2738\,X_3 \quad (21)$$

The Second Stage

Subtitute $\hat{y}_2$ into the equation (18) and $\hat{y}_1$ into equation (19) so that it is obtained:

$$Y_1^*=\gamma_{10}^*+\gamma_{12}^*\hat{y}_2+\beta_{11}^*X_1+\varepsilon_1^* \quad (22)$$
$$Y_2^*=\gamma_{20}^*+\gamma_{21}^*\hat{y}_1+\beta_{21}^*X_2+\beta_{23}X_3+\varepsilon_1^* \quad (23)$$

Apply OLS method to equation (22) and equation (23) to estimate the value of $\gamma_{10}^*, \gamma_{12}^*, \gamma_{20}^*\gamma_{21}^*, \beta_{11}^*, \beta_{21}^*$ dan β_{23}^*.

1. Apply OLS method to equation (22)

$$X^*=[1\quad \hat{y}_2\quad X_1]=\begin{bmatrix}1 & 131.82 & 7853\\1 & 132.80 & 7612\\1 & 132.17 & 7371\\\vdots & \vdots & \vdots\\1 & 126.19 & 11476\end{bmatrix}\; Y_1=\begin{bmatrix}135.72\\136.36\\136.34\\\vdots\\127.81\end{bmatrix}$$

$$\begin{bmatrix}\gamma_{10}\\\gamma_{12}\\\beta_{11}\end{bmatrix}=(X^{*\prime}X^*)^{-1}(X^{*\prime}Y_1)$$

2. Apply OLS method to equation (23)

$$X^*=[1\quad \hat{y}_1\quad X_2\quad X_3]=\begin{bmatrix}1 & 137.76 & 1344 & 123.07\\1 & 138.19 & 1363 & 123.36\\1 & 138.42 & 1355 & 123.59\\\vdots & \vdots & \vdots & \vdots\\1 & 128.37 & 131.17 & 115.44\end{bmatrix}\; Y_2=\begin{bmatrix}124.56\\1246\\125.49\\\vdots\\125.94\end{bmatrix}$$

$$\begin{bmatrix}\gamma_{20}\\\gamma_{21}\\\beta_{22}\\\beta_{23}\end{bmatrix}=(X^{*\prime}X^*)^{-1}(X^{*\prime}Y_2)$$

Based on the matlab program, the simultaneous equation model using 2SLS method:

$$\hat{Y}_1=-6.2470+1.14028\,Y_2-0.0008\,X_1$$
$$\hat{Y}_2=19.552+0.6483\,Y_1+0.3866\,X_2-0.2363X_3$$

4.3 Determination of k and Parameter Estimation Two Stage Ridge

Parameter value of $\boldsymbol{\alpha}_{2SLS}$ calculating with is

$$\boldsymbol{\alpha}_{2SLS} = \begin{bmatrix} -0.579076 \\ -0.326332 \\ -0.589659 \end{bmatrix}$$

While the MSE value obtained $\boldsymbol{\alpha}_{2SLS}$ is $\hat{\sigma}^2 = 0.091077$, then the itial of k value is:

$$k_1^0 = \frac{0.091077}{(-0.579076)^2} = 0.271606$$
$$k_2^0 = \frac{0.091077}{(-0.326332)^2} = 0.855242$$
$$k_3^0 = \frac{0.091077}{(-0.589659)^2} = 0.261943$$

After getting the value of $K = \text{diag}(k_1^0, k_2^0, k_3^0)$, The next step is to determine the parameter $\boldsymbol{\alpha}_{TR}$ through the iteration process. By using the help of Matlab software, the initial parameter of $(\hat{\boldsymbol{\alpha}}_{GR,j}^0)$ can be seen in the table 5.

Table 5. The Tesults of the Two Stage Ridge Parameter Estimator

Parameter	$\hat{\alpha}_{TR}^0$
$\hat{\alpha}_1$	-0.530634
$\hat{\alpha}_2$	-0.007606
$\hat{\alpha}_3$	-0.000314

To estimate parameter of two stage ridge, used paraeter value $\hat{\alpha}_{TR}^1$ in table 5 where $\hat{\beta}_{TR} = Q\hat{\alpha}_{TR}$ diperoleh

$$\hat{\beta}_{TR} = \begin{bmatrix} 0.306222 \\ 0.311508 \\ 0.301365 \end{bmatrix}$$

From the estimated parameter, the estimation results of the simutaneous equations model using two stage ridge are shown below:

$$Y_2 = 0.306222\, Y_1 + 0.311508\, X_2 + 0.301365\, X_3$$

To find out the best models of parameters in this study, namely by using bias and MSE values. A good estimator is an estimator who has the smallest bias and MSE value. From the output matlab obtained a ratio of variance, bias and mse values of each method.

Tabel 6. Variance, Bias and MSE

Method	Variance	Bias	MSE
β_{TR}	0.41839	0.00000	0.41839
β_{2SLS}	0.02831	0.45126	0.47957

Based on table 6, it can be seen that the two stage ridge method has more optimal bias and MSE values.

5. Conclusion

To find out the best model parameter estimator in this study is by using a comparison of the value of bias and MSE. A good estimator is an estimator that has the smallest bias and MSE values. From the result of the study, it was found that the two stage ridge is better than two stage least square. This is influenced by the simultaneous relationship and multicollinearity in each equation.

References

Koutsoyiannis, A., 1978. *Theory of Econometrics*, Bornes and Noble Books, New York, USA

Supramto, J., 1984, *Ekonometrik*. Penerbit Fakultas Ekonomi UI, Jakarta.

Syafaat, N., 1996, Pendugaan Parameter Persamaan Simultan dengan Metode Pendugaan OLS, 2SLS, LIML dan 3SLS, Ekonomi dan Keuangan Indonesia, 56 (4).

Shariff, N., Duzan, H., 2018. A Comparison of OLS and Ridge Regression Method in the Presence of Multicollinearity Problem in the Data, International Journal of Engineering & Technology, 7 (4.30), p. 36-38.

Eledum, H., Alkhalifa, A., 2012, Generalized Two Stages Ridge regression Estimator GTR for Multicollinearity and Autocorrelated Errors, Canadian Journal on Science and Engineering Mathematics,3(3).

Theil, H., 1953, *Repeated Least-Squares Applied to Complete Equation Systems*, The Netherlands: The Hague The Central Planning Bureau.

Basmann, R. L.,1957, Generalized Classical Method of Linear Estimation of Coefficients in a Structural Equation, Econometrica, (25): 77–83.

Gujarati, D., N., 2010, *Basic Econometrics*, New York: Mc Graw-hill Companies, ISBN:9780071276252

Lopez-Espin, J., J., Vidal., A., M., Giménez, D., 2012, Two-stage least squares and indirect least squares algorithms for simultaneous equations models, Journal of Computational and Applied Mathematics, 236(15): p.3676-3684.

Miftahul, U., Alfian, H. A., & Dian, A., 2011, "Analysis of Simultaneous Equation Model (SEM) on Non normally Response used the Method of Reduce Rank Vector Generalized Linear Models (RR-VGLM)", Proceeding The 1st IBSC: Towards The ExtendedUse Of Basic Science For Enhancing Health, Environment, Energy And Biotechnology, p. 316-318

Hoerl, A., E., Kennard, R., W., 1970, Ridge regression: applications to nonorthogonal problem, Technometrics, 12(1), p. 69-78.

Hoerl, A., E., Kennard, R., W., 1970, Ridge regression: biased estima- tion for nonorthogonal problems, Technometrics, 12(1), pp. 55-67.

Duzan, H., Shariff , N., S., M., 2015, "Ridge regression for solving the multicollinearity problem: review of methods and models".,Journal of Applied Sciences, 15(3), p. 392-404.

Mansson, K., Shukur, G., Kibria, B., M.,G., 2010, A simulation study of some ridge regression estimators under different distributional assumptions, Communications in Statistics- Simulation and Compu- tation, 39(8), p. 1639-1670.