Biresponses Kernel Nonparametric Regression: Inflation and Economic Growth

Suparti¹*, Budi Warsito¹, Rukun Santoso¹, Hasbi Yasin¹, Rezzy Eko Caraka² and Sudargo³

¹Departement of Statistics, Faculty Science and Mathematics, Diponegoro University, Semarang, Indonesia; ²Laboratory of Hierarchical Likelihood, Research for Basic Sciences, College of Natural Sciences, Seoul National University, Seoul 08826, South Korea; ³Departemen of Mathematics, PGRI University, Semarang, Indonesia

Abstract: The relation between inflation and economic growth is interesting to observe. To maintain the inflation rate, two factors should be taken into account, namely keeping the economic pulse at its optimal rate and keeping people’s purchasing power from decreasing. Many factors influence the inflation and economic growth of a nation; one of which is the national bank interest rate. Since the data of inflation are closely related to economic growth, this study aims at modelling the data of inflation rate and economic growth of Central Java Province in Indonesia using bi-response kernel regression. Employing the data from the first trimester of 2007 up to those from the second trimester of 2019 which were processed using kernel Gauss, the best model to minimise the value of GCV was obtained with optimum h for inflation model amounting to 0.12 and 81.75 for economic growth model. The model performance was excellent because the MAPE out sample was less than 10%. The biresponses kernel model is better than the linear biresponses model in terms of GCV, MSE, R², and MAPE values.

Keywords: Inflation, economic growth, biresponses, kernel.

I. INTRODUCTION

Inflation is an upward trend or movement of general prices which occurs continuously from one period to the next period. The increasing prices of goods can result from the increasing prices of raw materials, lacks supply and infrastructure, and decrease in currency rate (Dartanto 2010). The type of inflation caused by the increasing price of raw materials burdens both producers and consumers since the increasing price of raw materials will result in more expensive goods which can impact the purchasing power of the community (Suparti et al. 2018a) (Dartanto and Otsubo 2013). If the community purchasing power declines, people will reduce their expenses which then will slow down the economic growth because the most significant contributing component of economic growth is domestic consumption (Martini et al. 2012) (Anggoro 2015). The price hike can also occur due to the rising demand. This type of price hike is suitable for producers. The economy will develop more rapidly if people’s purchasing power rises.

The relation between inflation and economic growth is interesting to investigate because a shallow inflation rate will suppress economic growth (Suparti et al. 2016). On the other hand, a very high level of inflation will reduce people purchasing power which in turn will also disrupt the economic cycle (Zarrouk et al. 2016). Consequently, to maintain the inflation rate, two factors are essential to take into account simultaneously, namely the inflation rate which can produce an optimal economic pulse and keep the community purchasing power stable (Ifeacho and Ngalawa 2014). Besides, economic growth is also crucially needed to realise the development targets, such as creating more employment, improving national output, promoting tax collection, alleviating poverty, reducing unemployment, and increasing social welfare. A low level of economic growth will interfere with the realisation of those targets. Inflation and economic growth are closely linked.

An analysis method that can be used to model two or more data is regression analysis. Basically, in regression analysis, there are two types of data, i.e. time series data and cross-section data. Time series data are the data of a subject which are observed at multiple time from one to another period (Caraka et al. 2019d), (Caraka et al. 2019c). Cross-section data are the data of several subjects that are taken only once and independently on each subject. The combinations of time series data and cross-section data create longitudinal data (Wu and Zhang 2006). Longitudinal data consists of the data from multiple observations of each subject at different time intervals. These data are correlated among the same subjects and independent between different subjects. Biresponses regression analysis is used for the regression analysis on either time series, cross-section, or longitudinal data which involves two correlated variables of responses (Fernandes et al. 2014).

The researcher conducted modeling of time series data using kernel model and local polynomial to model the inflation data of Indonesia (Suparti et al. 2018b) (Suparti et al. 2014) (Suparti et al. 2018c). The researcher also conducted modelling of longitudinal data in the seven groups of inflation expenses using local spline polynomial and the combination of local
polynomial and truncated spline together or hybrid. The research related to biresponses data was conducted by (Wulandari and Budiantara 2014) (Chamidah and Rifada 2016) (Nurdiani et al. 2018) analyzed the factors influencing the percentage of the poor and per capita expenses in East Java Province using spline biresponses nonparametric regression. (Chamidah and Rifada 2016) estimated the curve of child growth media based on biresponses local linear estimator.

Moreover, (Ampulembang et al. 2016) created a model of the prosperity indicator of Java Island using MARS biresponses. Besides, (Pratiwi 2017) and (Nurdiani et al. 2018) modeled the truncated spline in biresponses nonparametric regression. Most of the modeling of biresponses data conducted by the previous researchers used local spline and polynomial methods. Meanwhile, the kernel nonparametric method was slightly discussed by (Lestari et al. 2019).

However, it was not yet applied to real data modeling. Therefore, in the research using other funding than from APBN FSM Undip, the researcher carried out a biresponses data modeling of inflation and economic growth of Central Java Province using kernel method. This research employed interest rate as a predicting factor since the interest rate variable had a significant impact on inflation (Agusmianata et al. 2017), and interest rate positively and significantly influenced the economic growth of Indonesia (Fahrika 2016). Using multiple regression modeling, it was revealed that there were specific relations among inflation, interest rate, and Indonesia’s economic growth of the period 2005-2015. Indonesia’s economic growth has a strong relation to inflation and interest rate, while inflation has a weak relation to interest rate (Indriyani 2016). In this paper, we will discuss the step construction of biresponses kernel nonparametric regression to the modeling of inflation and economic growth of central java.

II. METHODS

The idea in the approach through nonparametric regression is smoothing the curve (Eubank 2004). It means through the nonparametric regression approach assumes that the regression curve that is formed from the data is smooth (Caraka et al. 2019a) (Qin et al. 2016). The purpose of smoothing in regression is the existence of a specific function so that nonparametric regression has flexibility (de Boor 2002) (Marbun et al. 2020). Besides, we can minimize the variance of data and estimate the behaviour of data that tends to be different and have no influence so that the characteristics of the data (Farin 2002) (Ruppert et al. 2009) (Yandell and Eubank 1989). A pair of data \((X_i, Y_{(k)})\), \(k = 1,2, \ldots, n\) are given and inserted into biresponses regression model as follows:

\[
Y_{(k)} = g_{(k)}(X_i) + \epsilon_{(k)}, k = 1,2 \text{ and } a < X_i < b
\]

Where \(k\) is the response number. The regression functions \(g_{(1)}\) and \(g_{(2)}\) are not known and assumed to be unproblematic. \(\epsilon_{(k)}\) is an independently random error with zero mean and variant \(\sigma_{(k)}^2\) (Lestari et al. 2019). The objective of this biresponses nonparametric regression modelling is to find the estimation of \(g_{(k)}\) function whose format is not yet known (Suparti et al. 2019) (Devi et al. 2019). There are several approaches to estimate \(g_{(k)}\) Function, such as kernel method, local polynomial method, etc. One response either one predictor data using kernel method, which is model (1) with \(k = 1\), was already applied by the researcher to model the inflation data of Central Java Province (Suparti 2013). The one response regression estimator \(Y\) and one predictor \(X\) using \(K\) kernel with bandwidth \(h\) are known as Nadaraya Watson estimator which can be written as follows.

\[
g_{(k)}(x) = \frac{\sum_{i=1}^{n} K_h(x - X_i)Y_i}{\sum_{i=1}^{n} K_h(x - X_i)}
\]

\[
g_{(h)}(x) = \frac{\sum_{i=1}^{n} K_h(x - X_i)Y_i}{\sum_{i=1}^{n} K_h(x - X_i)}
\]

\[
g_{(h)}(x) = \frac{\sum_{i=1}^{n} K_h(X_i - X_j)Y_j}{\sum_{m=1}^{n} K_h(X_i - X_m)} = \sum_{j=1}^{n} W_{ij}Y_j
\]

with

\[
W_{ij} = \frac{K_h[X_i - X_j]}{\sum_{m=1}^{n} K_h[X_i - X_m]}
\]

Where \(K\) kernel function is a continuous, limited, symmetrical density function and \(K_h(x) = K(\frac{x}{h})\). So, Referring to Nadaraya Watson estimator \((2)\), kernel regression estimator at response-\(k\), \(g(k)\) with \(K\) kernel function, and bandwidth \(hk\) is as below:

\[
g_{(k)hk}(X_i) = \frac{\sum_{j=1}^{n} K_{hk}(X_i - X_j)X_{k(j)}}{\sum_{m=1}^{n} K_{hk}(X_i - X_m)} = \sum_{j=1}^{n} W_{(k)ij}Y_{(k)}^j, k = 1,2, \ldots
\]

\[
= (W_{(k)1}, W_{(k)2}, \ldots, W_{(k)n})^T \begin{pmatrix} Y_{(k)1} \\ Y_{(k)2} \\ \vdots \\ Y_{(k)n} \end{pmatrix} = W_{(k)}Y_{(k)}
\]

where \(Y_{(k)} = (Y_{(k)1}, Y_{(k)2}, \ldots, Y_{(k)n})^T\). Besides, \((W_{(k)1}, W_{(k)2}, \ldots, W_{(k)n})\) and \(W_{(k)ij} = \frac{K_{hk}[X_i - X_j]}{\sum_{m=1}^{n} K_{hk}[X_i - X_m]}\) if the data of response 1 are

\[
Y_{(1)} = \begin{pmatrix} Y_{(1)1} \\ Y_{(1)2} \\ \vdots \\ Y_{(1)n} \end{pmatrix}, \text{ the data of response 2 are } Y_{(2)} = \begin{pmatrix} Y_{(2)1} \\ Y_{(2)2} \\ \vdots \\ Y_{(2)n} \end{pmatrix}
\]
were tri-monthly data of 2007-2019. The first tri-monthly data were recorded in March; and the second, third, and fourth ones were in June, September, and December respectively. Also, the economic growth data were calculated from GRDP growth rate of Central Java (year on year). The inflation data used were the inflation data of Central Java Province (yoy) in the same months, i.e. March, June, September, and December. Also, the interest rates of BI used were the data from March, June, October, and December during the period. The data analysis procedure included: 1) Doing a theoretical study on biresponses regression model using kernel method; 2) Dividing the data into two categories, namely in-sample data and out sample data; 3) Modeling in-sample data by making prior identification of the kernel function, then followed by determining the different sizes of bandwidth combinations for different responses of model 1 and model 2. Next, the value of GCV for each combination of bandwidth was calculated, and the minimum value of GCV was selected to determine the best model; 4) Evaluating the performances of the models by calculating the MAPE out a sample; 5) Concluding.

A. The Modelling of Inflation and Economic Growth Data Based on the Interest Rate of the Bank of Indonesia

In this case, response $Y_1$ variable was inflation; response $Y_2$ variable was economic growth, and predictor X variable was an interest rate of BI. The data were divided into two, namely in-sample data from the first trimester of 2007 up to the fourth trimester of 2016 and out sample data starting from the first trimester of 2017 up to the second trimester of 2019. The in-sample data were used to make a model, and the out sample data were to evaluate the model performance based on MAPE(Caraka et al. 2019d). MAPE can be formulated as follows:

$$\text{MAPE} = \frac{1}{N} \sum_{k=1}^{2} \sum_{i=1}^{n} \left| \frac{Y_{(k)i} - \hat{Y}_{(k)i}}{Y_{(k)i}} \right| \times 100\%$$

Where:
$Y_{(k)i}$ : the actual value of data on $i^{th}$ observation on $k^{th}$ response
$\hat{Y}_{(k)i}$ : predictive value of $Y$ on $i^{th}$ observation for $k^{th}$ response
$n$ : number of observations on each response
$N$ : $2n$

The smaller the MAPE value, the better the model performance. (Caraka et al. 2019b; Suhermi et al., 2018; Suhartono and Enderharto, 2009) argued that the evaluation values obtained would fall into the criteria as in Table 1.
Table 1: Mape Justification

<table>
<thead>
<tr>
<th>Measure</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;4.9%</td>
<td>High Accurate</td>
</tr>
<tr>
<td>5%-9.9%</td>
<td>Accurate Forecasting</td>
</tr>
<tr>
<td>10%-14.9%</td>
<td>Good Forecasting</td>
</tr>
<tr>
<td>15%-19.9%</td>
<td>Reasonable Forecasting</td>
</tr>
<tr>
<td>&gt;20%</td>
<td>Inaccurate Forecasting</td>
</tr>
</tbody>
</table>

The modelling was conducted for in-sample data on the first response and second response using kernel Gauss by varying the combination of bandwidth sizes. Then, GCV minimum value was obtained as much as 1.238515 with optimum bandwidth size $h_1 = 0.12$ and $h_2 = 81.75$ and $R^2 = 1.498503$. The scatter plots of the data and models are shown below in Figures 1 and 2.

Figure 1: Inflation model using kernel.

Figure 2: Economic growth model using kernel.

From the optimal model which was already obtained, the out sample MAPE value was calculated amounting 6.611294% which showed that the model performance was excellent because the MAPE value was < 10%. The prediction of out sample data is presented in Figures 3 and 4 below.

Figure 3: Prediction of out sample data on inflation using kernel model.

Figure 4: Prediction of outsample data on economic growth using kernel model.

B. Comparison between Biresponses Kernel Regression and Biresponses Linear Regression

The scatter plots of inflation data and economic growth data based on BI interest rates (Figures 1 and 2) showed the tendency of straight-line patterns so that they could be further processed using biresponses linear regression. Therefore, the researcher was interested in comparing the kernel model and biresponses linear model of the inflation and economic growth data of Central Java Province. If the linear model of (1) is

$$Y_{(k)i} = a_{(k)} + b_{(k)}X_i + e_{(k)i}, \ k = 1,2; i = 1,2,...,n$$

If the data on response 1 are $Y_{(1)} = \begin{pmatrix} Y_{(1)1} \\ Y_{(1)2} \\ \vdots \\ Y_{(1)n} \end{pmatrix}$, the data on response 2 are $Y_{(2)} = \begin{pmatrix} Y_{(2)1} \\ Y_{(2)2} \\ \vdots \\ Y_{(2)n} \end{pmatrix}$, and the data on predictor variable are $X_1 = \begin{pmatrix} 1 \\ X_1 \\ \vdots \\ X_n \end{pmatrix}$, parameters $\theta_1 = \begin{pmatrix} a_{(1)} \\ b_{(1)} \end{pmatrix}$, $\theta_2 = \begin{pmatrix} a_{(2)} \\ b_{(2)} \end{pmatrix}$ and errors $\varepsilon_1 = \begin{pmatrix} e_{(1)1} \\ e_{(1)2} \\ \vdots \\ e_{(1)n} \end{pmatrix}$, $\varepsilon_2 = \begin{pmatrix} e_{(2)1} \\ e_{(2)2} \\ \vdots \\ e_{(2)n} \end{pmatrix}$ then the biresponse linear regression model can be written in matrix as follows.

$$\begin{pmatrix} Y_{(1)} \\ Y_{(2)} \end{pmatrix} = \begin{pmatrix} X_1 & 0 \\ 0 & X_1 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}.$$
Or
\[ Y = X_1 \theta + \varepsilon \]

With: \( Y = \begin{bmatrix} Y_1(1) \\ Y_1(2) \\ \vdots \\ Y_1(n) \\ Y_2(1) \\ Y_2(2) \\ \vdots \\ Y_2(n) \end{bmatrix}_{2n \times 1} \)

\[ X_1 = \text{diag}(X_1, X_1) = \begin{bmatrix} X_1 & 0 \\ 0 & X_1 \end{bmatrix} = \begin{bmatrix} 1 & X_1 & 0 & 0 \\ 1 & X_2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_n & 0 & 0 \\ 0 & 0 & 1 & X_1 \\ 0 & 0 & 1 & X_2 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & X_n \end{bmatrix}_{(2n \times 4)} \]

\[ \theta = \begin{bmatrix} \theta_{(1)} \\ \theta_{(2)} \end{bmatrix} = \begin{bmatrix} \theta_{(1)1} \\ \theta_{(1)2} \\ \vdots \\ \theta_{(1)n} \\ \theta_{(2)1} \\ \theta_{(2)2} \\ \vdots \\ \theta_{(2)n} \end{bmatrix}_{4 \times 1} \] and \( \varepsilon = \begin{bmatrix} \varepsilon_1(1) \\ \varepsilon_1(2) \\ \vdots \\ \varepsilon_1(n) \\ \varepsilon_2(1) \\ \varepsilon_2(2) \\ \vdots \\ \varepsilon_2(n) \end{bmatrix}_{4 \times 1} \).

The estimator of linear model was completed using ordinary least square by minimizing the sum of the squares of errors, \( \varepsilon^T \varepsilon = (Y - X' \theta)^T (Y - X' \theta) \) and the result obtained was \( \hat{\theta} = (X^{-1}X')^{-1}X^{-1}Y \). Using R software computation (R Core Team 2008), the optimum linear model was produced with the values of MSE= 1.503206, GCV= 1.665602 dan \( R^2 = 0.500679 \).

The linear models formed are presented in Figures 5 and 6.

From the linear models generated, the MAPE value of out sample was counted. The result was 32.28913% which indicated that the model performance was reasonably good since the value was 20% \( \leq \text{MAPE} \leq 50% \). The graphics of out sample data are demonstrated in Figures 7 and 8.

After performing the models, we can evaluate the accuracy of these models by MSE, \( R^2 \) and MAPE out values of the sample which present in Table 2.

Based on Table 2, the two models applied on the inflation and economic growth data of Central Java Province, namely biresponse kernel model and biresponse linear model revealed that the values of MSE and GCV in biresponse kernel model were smaller, and the value of \( R^2 \) in biresponse linear model was larger. From the performance evaluation, based on the MAPE value of out sample, biresponse kernel model showed an outstanding performance because the MAPE value was < 10%.

IV. CONCLUSION

The modelling of the inflation and economic growth data of Central Java Province using biresponse kernel...
model which involved one predictor variable namely the interest rates of the Bank of Indonesia showed an excellent performance. Even though the scatter plots of the data indicated similar patterns to those of the linear model, the approach using biresponses linear regression did not demonstrate a better performance compared to the findings of the kernel model. In further research, inflation and economic growth data modelling can be conducted using biresponses kernel and local polynomial regression.

ACKNOWLEDGMENT

This work is supported by the Faculty of Science and Mathematics, Diponegoro University under contract 4858/UN7.5.8/PP/2019.

REFERENCES


Anggoro Y (2015) Automotive and Logistics Clusters To Drive Economic. The University of North Carolina at Charlotte


de Boor C (2002) Chapter 6 - Spline Basics


Fahrika A (2016) Pengaruh tingkat suku bunga melalui investasi swasta terhadap pertumbuhan. Econ Soc Dev Studies 3:


Biresponses Kernel Nonparametric Regression

International Journal of Criminology and Sociology, 2021, Vol. 10


DOI: https://doi.org/10.6000/1929-4409.2021.10.54

© 2021 Suparti et al.; Licensee Lifescience Global. This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.