Quest Journals
Journal of Research in Applied Mathematics
Volume 11 ~ Issue 2 (2025) pp: 53-61
ISSN (Online): 2394-0743 ISSN (Print): 2394-0735
www.questjournals.org

Review Paper

Modeling and Dependence among Factors Based on Multivariate Time Series Data

Ying Han[1], Ding Peng[2], Mingduan Liang[3]
[1,2,3]School of Mathematical Sciences, Guizhou Normal University, Guiyang, China
Corresponding Author: Ying Han

ABSTRACT: *To investigate the relationship between personal consumption expenditure and personal disposable income, this paper combines them into multivariate time series data and conducts a detailed discussion on the model of this data. With multiple candidate models obtained, the best model is finally selected by analyzing their goodness of fit and autocorrelation. Furthermore, the causal relationship, namely the dynamic dependence, between personal consumption expenditure and personal disposable income is explored.*

KEYWORDS: *VARMA, Granger Causality, Maximum Likelihood Estimation*

R*eceived 08 Feb., 2025;* ***R****evised 16 Feb., 2025;* ***A****ccepted 18 Feb., 2025*

I. INTRODUCTION

Time series data analysis has been around for quite some time, but the real research began with the AR model proposed by Yule [1]. After its introduction, it has been widely applied across multiple disciplines. Walker [2] proposed the MA model and the ARMA model, which have since become the foundation of time series analysis. Subsequently, Box and Jenkins [3] introduced the ARIMA model, which, together with the AR, MA, and ARMA models, forms the basis for the study of stationary linear models. For identification and estimation of AR, MA, and ARMA models in the context of univariate data, references [4]-[10] can be consulted.

Regarding the research on multivariate time series, it is typically an extension of the methods used for univariate time series models, such as the vector autoregressive (VAR) model, VARMA model, and so on. Sims [11] proposed the relevant theory of the VAR model, which is an extension of the AR model. Under specific conditions, multivariate MA and ARMA models can also be transformed into VAR models. Therefore, it is often used for the estimation and prediction of multiple related variables or indicators. However, the VAR model also has significant drawbacks: first, as shown in the article by Lütkepohl and Poskitt [12], they are generally not as concise as VARMA models. Second, Lütkepohl [13] proved in 1999 that the VAR model family is not closed under marginalization and temporal aggregation. Tiao and Box [14] proposed methods for modeling multivariate time series analysis and studied the properties of VARMA models. For the non-stationary multivariate case, Granger [15] proposed cointegration theory to characterize the long-term change relationships between sequences. Hendry and Anderson [16] provided the ECM (Error Correction) model to describe the short-term change relationships between sequences.

Theoretical research on VARMA models can be found in [17] and [18]. What has attracted the interest of many researchers is how to estimate the parameters, especially the parameter estimates obtained by maximizing the nonlinear likelihood function, which still lack analytical forms to date. Whittle [19] and Durbin [20] developed approximations for the univariate MA model. Godolphin [21] extended Whittle's method to the univariate ARMA model and showed how to compute parameter estimates directly from the sample autocorrelations of the observed data. Durbin's approximation leads to an algorithm where the MA coefficients are calculated from the coefficients of a long autoregression. Hannan and Rissanen [22] and Koreisha and Pukkila [23] extended the Durbin method to univariate and multivariate ARMA models, respectively. Similar approximate MLE methods have been proposed by Tunnicliffe-Wilson [24], Reinsel et al. [25], and DeFrutos and Serrano [26].

The structure of this paper is as follows: Section 2 introduces the VARMA model. Section 3 discusses the methods for identifying dynamic dependence, namely Granger causality. Section 4 estimates the model parameters using the conditional maximum likelihood method. Section 5 examines the relationship between personal consumption expenditure and personal disposable income.

II. VARMA MODEL

The definition equation of the VARMA(p,q) model is:

$$z_t = \phi_0 + \sum_{i=1}^{p} \phi_i z_{t-i} + a_t - \sum_{j=1}^{q} \theta_q a_{t-j} \qquad (1)$$

where z_t is a $k \times 1$ dimensional time series data, ϕ_0 is a k-dimensional constant vector, ϕ_i and θ_i (for $i \geq 1$) are $k \times k$ matrices, and a_t is a sequence of independently and identically distributed random vectors with mean 0 and covariance matrix Σ_a, which is a positive definite matrix. Using a compact notation, Equation (1) can be rewritten as follows:

$$\phi(B) z_t = \phi_0 + \theta(B) a_t \qquad (2)$$

Where $\phi(B) = I_k - \phi_1 B - \cdots - \phi_p B^p$ and $\theta(B) = I_k - \theta_1 B - \cdots - \theta_q B^q$, The backshift operator B is defined as:

$$B z_t = z_{t-1} \qquad (3)$$

Next, we will present several theorems that are useful for the subsequent formula derivation and application. However, we will not provide the proofs of these theorems. For detailed proofs, please refer to Reference [27].

Theorem 2.1: For the model in Equation (2), the sufficient conditions for the identifiability of the VARMA model are: 1. $\phi(B)$ and $\theta(B)$ are left coprime. 2.The order of the MA part *q* and the order of the AR part *p* should be as small as possible, while the matrices $\phi_p(p) > 0$ and $\theta_q(q) > 0$ satisfy that the rank of the combined matrix $[\phi_p, \theta_q]$ is k, where k is the dimension of z_t.

Note: Condition 1 means that if $u(B)$ is a common left factor of $\phi(B)$ and $\theta(B)$,then $|u(B)|$ is a nonzero constant, and the polynomial matrix $u(B)$ is called a unimodular matrix. For more discussion on Condition 2, please refer to [28] and [29].

The identifiability of VARMA models indicates that the specification of these models involves not only the identification of the orders (p,q) but also the structure.

Theorem 2.2: The process z_t in Equation (2) is weakly stationary if and only if all the roots of the determinant $|\phi(B)| = 0$ lie outside the unit circle, meaning that their absolute values are all greater than 1.

Theorem 2.3: For the model in Equation (2), the necessary and sufficient condition for the invertibility of z_t is that all the roots of the determinant $|\theta(B)| = 0$ lie outside the unit circle, meaning that their absolute values are all greater than 1.

III. DYNAMIC DEPENDENCE

Granger causality can be used to determine whether there is dynamic dependence between a component at the current time and all components at the previous time or several previous times, that is, whether the information from the previous time or several previous times improves the prediction of a component at the current time. The MA representation of the VARMA model is used to judge the dynamic relationships between variables. When z_t is stationary, the MA representation can be written as:

$$z_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots = \psi(B) a_t \qquad (4)$$

Among them, the definition of $\psi(B)$ is given by: $\psi(B) = \sum_{i=0}^{\infty} \psi_i B^i$. The unified equation for ψ_i is:

$$\psi_i = \sum_{j=1}^{\min\{p,i\}} \phi_i \psi_{i-j} - \theta_j \qquad (5)$$

In particular, $\theta_i = 0$ for $i > q$ and $\psi_0 = 0$. Considering the simple partition: $z_t = (z'_{1t}, z'_{2t})'$ and $a_t = (u'_{1t}, u'_{2t})'$, Equation (4) can be rewritten as:

$$\begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix} = \begin{pmatrix} \psi_{11}(B) & \psi_{12}(B) \\ \psi_{21}(B) & \psi_{22}(B) \end{pmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \tag{6}$$

It should be noted that if $\psi_{12}(B) = 0$, but $\psi_{21}(B) \neq 0$, then there exists a unidirectional relationship from z_{1t} to z_{2t}. z_{1t} does not depend on any past information of z_{2t}, but z_{2t} depends on some lagged values of z_{1t}. This indicates that z_{1t} is a cause of z_{2t}, and it also suggests the existence of a linear transfer function model within the VARMA model. Similarly, the VARMA model in Equation (2) can also be rewritten after mean adjustment as:

$$\begin{pmatrix} \phi_{11}(B) & \phi_{12}(B) \\ \phi_{21}(B) & \phi_{22}(B) \end{pmatrix} \begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix} = \begin{pmatrix} \theta_{11}(B) & \theta_{12}(B) \\ \theta_{21}(B) & \theta_{22}(B) \end{pmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \tag{7}$$

Using the definition of $\psi(B)$, we have:

$$\begin{pmatrix} \phi_{11}(B) & \phi_{12}(B) \\ \phi_{21}(B) & \phi_{22}(B) \end{pmatrix}^{-1} \begin{pmatrix} \theta_{11}(B) & \theta_{12}(B) \\ \theta_{21}(B) & \theta_{22}(B) \end{pmatrix} = \begin{pmatrix} \psi_{11}(B) & \psi_{12}(B) \\ \psi_{21}(B) & \psi_{22}(B) \end{pmatrix} \tag{8}$$

Thus, we obtain the following equation:

$$[\phi(B)]^{-1} = \begin{pmatrix} D(B) & -D(B)\phi_{12}(B)\phi_{22}^{-1}(B) \\ -\phi_{22}{}^{-1}(B)\phi_{21}(B)D(B) & \phi_{22}{}^{-1}(B) + \phi_{22}{}^{-1}(B)\phi_{21}(B)D(B)\phi_{12}(B)\phi_{22}^{-1}(B) \end{pmatrix} \tag{9}$$

where $D(B)$ is defined as follows:

$$D(B) = (\phi_{11}(B) - \phi_{12}(B)\phi_{22}^{-1}(B)\phi_{21}(B))^{-1} \tag{10}$$

By substituting Equation (9) into Equation (8), we can obtain that when $\psi_{12}(B) = 0$, it is equivalent to

$$D(B)\,[\theta_{12}(B) - \phi_{12}(B)\phi_{22}^{-1}(B)\theta_{22}(B)] = 0 \tag{11}$$

When $\psi_{21}(B) \neq 0$, it is equivalent to

$$\theta_{21}(B) - \phi_{21}(B)\phi_{11}^{-1}(B)\theta_{11}(B) \neq 0 \tag{12}$$

Therefore, if Equations (11) and (12) hold, it indicates that there is a Granger causality from z_{1t} to z_{2t}.

IV. MODEL ESTIMATION

Before proceeding with the calculations, we first apply the three theorems mentioned in Section 2. The VARMA model must satisfy these three theorems—identifiability, stationarity, and invertibility—in order to proceed with the subsequent derivations. We assume that the sequence values before all sampling times are zero, i.e., $a_t = 0$ and $z_t = \bar{z}$ for $t \leq 0$, where $\bar{z}$ is the sample mean, and $E(z_t) = 0$. Define $\boldsymbol{Z} = \{z'_1, z'_2, \cdots, z'_t\}'$ and $\boldsymbol{A} = \{a'_1, a'_2, \cdots, a'_T\}'$, for $t = 1, 2, \cdots, T$. The general VARMA model in Equation (1) can be rewritten as:

$$\boldsymbol{\Phi Z} = \boldsymbol{\Theta A} \tag{13}$$

Where

$$\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{I_k} & 0_{\boldsymbol{k}} & 0_{\boldsymbol{k}} & \cdots & 0_{\boldsymbol{k}} \\ -\boldsymbol{\phi}_1 & \boldsymbol{I_k} & 0_{\boldsymbol{k}} & \cdots & 0_{\boldsymbol{k}} \\ -\boldsymbol{\phi}_2 & -\boldsymbol{\phi}_1 & \boldsymbol{I_k} & \cdots & 0_{\boldsymbol{k}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_{\boldsymbol{k}} & 0_{\boldsymbol{k}} & 0_{\boldsymbol{k}} & \cdots & \boldsymbol{I_k} \end{bmatrix}, \boldsymbol{\Theta} = \begin{bmatrix} \boldsymbol{I_k} & 0_{\boldsymbol{k}} & 0_{\boldsymbol{k}} & \cdots & 0_{\boldsymbol{k}} \\ -\boldsymbol{\theta}_1 & \boldsymbol{I_k} & 0_{\boldsymbol{k}} & \cdots & 0_{\boldsymbol{k}} \\ -\boldsymbol{\theta}_2 & -\boldsymbol{\theta}_1 & \boldsymbol{I_k} & \cdots & 0_{\boldsymbol{k}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_{\boldsymbol{k}} & 0_{\boldsymbol{k}} & 0_{\boldsymbol{k}} & \cdots & \boldsymbol{I_k} \end{bmatrix} \tag{14}$$

Both $\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}$ are $kT \times kT$ matrices, 0_k is a $k \times k$ zero matrix, and $|\boldsymbol{\Theta}| = 1$. Let $\boldsymbol{W} = \boldsymbol{\Phi Z}$, then $\boldsymbol{A} = \boldsymbol{\Theta}^{-1}\boldsymbol{W}$. The transformation from $\boldsymbol{A}$ to $\boldsymbol{W}$ has a unit Jacobian. Under the assumption of normality, $\boldsymbol{A} \sim N(0, \boldsymbol{I_t} \otimes \boldsymbol{\Sigma_a})$. Therefore, the conditional log-likelihood function of the data is:

$$l(\boldsymbol{\beta}, \boldsymbol{\Sigma_a}; \boldsymbol{Z}) = -\frac{T}{2}\log(|\boldsymbol{\Sigma_a}|) - \frac{1}{2}\boldsymbol{A}'(\boldsymbol{I_T} \otimes \boldsymbol{\Sigma_a^{-1}})\boldsymbol{A} \tag{15}$$

Where $\boldsymbol{\beta} = vec[\boldsymbol{\phi}_1, \cdots, \boldsymbol{\phi_p}, \boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta_q}]$, Taking the derivative of Equation (15) yields:

$$\frac{\partial l(\boldsymbol{\beta}, \boldsymbol{\Sigma_a}; \boldsymbol{Z})}{\partial \boldsymbol{\Sigma_a}} = -\frac{T}{2}\boldsymbol{\Sigma_a^{-1}} + \frac{1}{2}\sum_{t=1}^{n}(\boldsymbol{\Sigma_a^{-1}}\boldsymbol{a_t}\boldsymbol{a_t}^{-1}\boldsymbol{\Sigma_a^{-1}}) \tag{16}$$

Therefore, given $\boldsymbol{\beta}$, the conditional MLE (Maximum Likelihood Estimation) for $\boldsymbol{\Sigma_a}$ is: $\widehat{\boldsymbol{\Sigma_a}} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{a_t}\boldsymbol{a_t}^{-1}$. Since $\boldsymbol{\theta_i}\boldsymbol{a_{t-i}} = (\boldsymbol{a_{t-i}'}\} \otimes \boldsymbol{I_k})vec(\boldsymbol{\theta_i})$, combining this with the general VARMA model in Equation (1) yields: at=

$$\boldsymbol{a_t} = \boldsymbol{z_t} - \sum_{i=1}^{p}(\boldsymbol{z_{t-i}'} \otimes \boldsymbol{I_k})vec(\boldsymbol{\phi_i}) + \sum_{j=1}^{q}\boldsymbol{\theta_q}\boldsymbol{a_{t-j}} \tag{17}$$

Taking the expectation first and then the partial derivative of Equation (16) yields:

$$\frac{\partial \boldsymbol{a_t'}}{\partial vec(\boldsymbol{\phi_i})} = \sum_{j=1}^{q}\frac{\partial \boldsymbol{a_{t-j}'}}{\partial vec(\phi_i)}\boldsymbol{\theta_j'} - (z_{t-i} \otimes \boldsymbol{I_k}), i = 1, 2, \cdots, p \tag{18}$$

$$\frac{\partial \boldsymbol{a_t'}}{\partial vec(\boldsymbol{\theta_j})} = \sum_{i=1}^{q}\frac{\partial \boldsymbol{a_{t-i}'}}{\partial vec(\boldsymbol{\theta_j})}\boldsymbol{\theta_i'}\} + (\boldsymbol{a_{t-j}} \otimes \boldsymbol{I_k}), j = 1, 2, \cdots, q \tag{19}$$

Define $u_{i,t} = \frac{\partial \boldsymbol{a_t'}}{\partial vec(\boldsymbol{\phi_i})}$ and $v_{j,t} = \frac{\partial \boldsymbol{a_t'}}{\partial vec(\boldsymbol{\theta_j})}$, then Equations (18) and (19) can be transformed into:

$$u_{i,t} - \sum_{j=1}^{p}u_{i,t-j}\boldsymbol{\theta_j'}\} = -(z_{t-i} \otimes \boldsymbol{I_k}) \tag{20}$$

$$v_{j,t} - \sum_{i=1}^{q}v_{j,t-i}\boldsymbol{\theta_i'}\} = +(\boldsymbol{a_{t-j}} \otimes \boldsymbol{I_k}) \tag{21}$$

From the log-likelihood function of Equation (15), we can derive that:

$$\frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \frac{\partial \boldsymbol{A}'}{\partial \boldsymbol{\beta}(\boldsymbol{I_T} \otimes \boldsymbol{\Sigma_a^{-1}})\boldsymbol{A}} \tag{22}$$

Therefore, the partial derivative of $-\boldsymbol{a_t'}$ with respect to $\boldsymbol{\beta}$ can be calculated and denoted as m_t:

$$\frac{-\partial \boldsymbol{a_t'}}{\partial vec(\boldsymbol{\beta})} = -\begin{pmatrix} u_{1,t} \\ \vdots \\ u_{p,t} \\ v_{1,t} \\ \vdots \\ v_{q,t} \end{pmatrix} \equiv m_t \tag{23}$$

It is not difficult to see that m_t is a $k^2(p+q) \times k$ matrix. Therefore, the matrix form of the partial derivative of $-\boldsymbol{A}'$ with respect to $\boldsymbol{\beta}$ is $M = (m_1, \cdots, m_T)$, where M is a $k^2(p+q) \times kT$ matrix. Combining Equations (20) and (21), the matrix n_t is defined as follows:

$$n_t = \begin{pmatrix} z_{t-1} \otimes \boldsymbol{I_k} \\ \vdots \\ z_{t-p} \otimes \boldsymbol{I_k} \\ -\boldsymbol{a_{t-1}} \otimes \boldsymbol{I_k} \\ \vdots \\ \boldsymbol{a_{t-q}} \otimes \boldsymbol{I_k} \end{pmatrix}, t = 1, 2, \cdots, T \tag{24}$$

then

$$M \times \begin{bmatrix} \boldsymbol{I_k} & -\boldsymbol{\theta}_1 & -\boldsymbol{\theta}_2 & \cdots & 0_{\boldsymbol{k}} \\ 0_{\boldsymbol{k}} & \boldsymbol{I_k} & -\boldsymbol{\theta}_1 & \cdots & 0_{\boldsymbol{k}} \\ 0_{\boldsymbol{k}} & 0_{\boldsymbol{k}} & \boldsymbol{I_k} & \cdots & 0_{\boldsymbol{k}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_{\boldsymbol{k}} & 0_{\boldsymbol{k}} & 0_{\boldsymbol{k}} & \cdots & \boldsymbol{I_k} \end{bmatrix} = M \times \boldsymbol{\Theta}' = \boldsymbol{N} \tag{25}$$

Where $\boldsymbol{N} = (n_1, n_2, \cdots, n_T)$. Therefore, the following partial derivatives can be obtained:

$$\frac{-\partial \boldsymbol{A}'}{\partial \boldsymbol{\beta}} = M = N(\boldsymbol{\Theta}')^{-1} \tag{26}$$

By combining Equations (22) and (26), we have:

$$\frac{-\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \quad N(\boldsymbol{\Theta}')^{-1}(\boldsymbol{I_T} \otimes \boldsymbol{\Sigma}_{\boldsymbol{a}}^{-1})\boldsymbol{A} \tag{27}$$

Setting the above equation to zero yields the normal equations for the log-likelihood function of a stationary and invertible VARMA model. When divided by $T\,(T \to \infty)$, the neglected terms converge in probability to zero. The Hessian matrix of the log-likelihood function is:

$$\frac{\partial^2}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} \approx \frac{\partial \boldsymbol{A}'}{\partial \boldsymbol{\beta}} (\boldsymbol{I_T} \otimes \boldsymbol{\Sigma}_{\boldsymbol{a}}^{-1}) \frac{\partial \boldsymbol{A}}{\partial \boldsymbol{\beta}'} \approx N(\boldsymbol{\Theta}')^{-1}(\boldsymbol{I_T} \otimes \boldsymbol{\Sigma}_{\boldsymbol{a}}^{-1})\boldsymbol{\Theta}^{-1} N' \tag{28}$$

V. EMPIRICAL ANALYSIS

The data used in this section, covering 639 observations of personal consumption expenditure (PCE) and personal disposable income (DSPI) from January 1959 to 2012, can be accessed via the following link: PCE AND DSPI.

The raw data are bivariate, with z_1 denoting personal consumption expenditure and z_2 denoting personal disposable income. Since the raw data are not stationary, this paper takes the logarithm of the data and then applies differencing to achieve stationarity.

p \ q	0	1	2	3	4	5	6
0	00000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0120
1	0.0000	0.0005	0.0003	0.0874	0.2523	0.2738	0.7914
2	0.0000	0.0043	0.0054	0.9390	0.4237	0.3402	0.8482
3	0.0000	0.8328	0.9397	0.9965	0.9376	0.9100	0.8193
4	0.0003	0.9643	0.9797	0.9937	0.9701	0.9810	0.9620
5	0.0150	1.0000	1.0000	1.0000	0.9995	0.9997	0.9851
6	0.1514	1.0000	1.0000	1.0000	1.0000	1.0000	0.9985

Table 5-1: p-values of the Extended Cross-Correlation Matrix

Table 5-1 presents the p-values of the extended cross-correlation matrix for the differenced logarithmic data. The maximum autoregressive order and the maximum moving average order are both set to 6. The rows and columns of the table represent the changes in p-values as the autoregressive and moving average orders increase from 0 to 6, respectively. It should be noted that these orders are different from the autoregressive order p. As the AR order increases, the p-values gradually increase, indicating a gradual decrease in the autocorrelation of the residual series. For the MA order, when the AR order is fixed, increasing the MA order typically leads to an increase in p-values, indicating a decrease in the autocorrelation of the residual series. Therefore, the autoregressive order of the VARMA model may be 2, 3, or 4, and the moving average order may be 1 or 2. Further experiments will be conducted to select the best model.

	(2,1)	(3,1)	(4,1)	(2,2)	(3,2)	(4,2)
AIC	-2.086201	-2.102131	-1.998464	-2.027707	-2.081865	-1.002797
BIC	-2.002346	-1.990323	-1.928584	-1.908911	-1.942105	-0.821109

Table 5-2: Goodness-of-Fit Test

From Table 5-2, it can be seen that the AIC and BIC values for VARMA(2,1) and VARMA(3,1) are significantly lower than those of the other four models. Lower AIC and BIC values indicate better model fit. Specifically, VARMA(2,1) has a lower AIC value than VARMA(3,1), but VARMA(3,1) has a lower BIC value. However, this alone does not definitively determine which model is better, hence further analysis is conducted to select the optimal model.

```
Ljung-Box Statistics:
           m       Q(m)      df    p-value
 [1,]   1.000     0.268    4.000     0.99
 [2,]   2.000     4.914    8.000     0.77
 [3,]   3.000    15.112   12.000     0.24
 [4,]   4.000    21.279   16.000     0.17
 [5,]   5.000    24.317   20.000     0.23
 [6,]   6.000    35.496   24.000     0.06
 [7,]   7.000    38.452   28.000     0.09
 [8,]   8.000    39.831   32.000     0.16
 [9,]   9.000    41.917   36.000     0.23
[10,]  10.000    45.199   40.000     0.26
[11,]  11.000    51.557   44.000     0.20
[12,]  12.000    57.162   48.000     0.17
[13,]  13.000    62.896   52.000     0.14
[14,]  14.000    64.337   56.000     0.21
[15,]  15.000    69.027   60.000     0.20
[16,]  16.000    74.269   64.000     0.18
[17,]  17.000    77.470   68.000     0.20
[18,]  18.000    79.330   72.000     0.26
[19,]  19.000    87.516   76.000     0.17
[20,]  20.000    91.887   80.000     0.17
[21,]  21.000    97.850   84.000     0.14
[22,]  22.000   101.108   88.000     0.16
[23,]  23.000   106.065   92.000     0.15
[24,]  24.000   118.724   96.000     0.06
Hit Enter to obtain residual plots:
```

Figure 5-1: The autocorrelation tests for VARMA(2,1)

```
Ljung-Box Statistics:
          m      Q(m)     df    p-value
 [1,]   1.000    0.244   4.000    0.99
 [2,]   2.000    0.305   8.000    1.00
 [3,]   3.000    0.838  12.000    1.00
 [4,]   4.000    5.163  16.000    0.99
 [5,]   5.000    6.191  20.000    1.00
 [6,]   6.000   14.968  24.000    0.92
 [7,]   7.000   18.204  28.000    0.92
 [8,]   8.000   19.370  32.000    0.96
 [9,]   9.000   20.850  36.000    0.98
[10,]  10.000   24.729  40.000    0.97
[11,]  11.000   30.537  44.000    0.94
[12,]  12.000   36.901  48.000    0.88
[13,]  13.000   44.310  52.000    0.77
[14,]  14.000   46.160  56.000    0.82
[15,]  15.000   51.098  60.000    0.79
[16,]  16.000   56.487  64.000    0.74
[17,]  17.000   59.024  68.000    0.77
[18,]  18.000   60.618  72.000    0.83
[19,]  19.000   68.151  76.000    0.73
[20,]  20.000   72.235  80.000    0.72
[21,]  21.000   79.746  84.000    0.61
[22,]  22.000   81.338  88.000    0.68
[23,]  23.000   86.954  92.000    0.63
[24,]  24.000  101.260  96.000    0.34
Hit Enter to obtain residual plots:
```

Figure 5-2: The autocorrelation tests for VARMA(3,1).

From the above two figures, it can be seen that although the p-values of the Ljung-Box test for both models are greater than 0.05, indicating that the null hypothesis of no significant autocorrelation in the residuals cannot be rejected, Figure 5.2 shows that almost all p-values are higher than those in Figure 5.1. This suggests that the autocorrelation in the residuals of the VARMA(3,1) model is even less significant, indicating a better model fit.

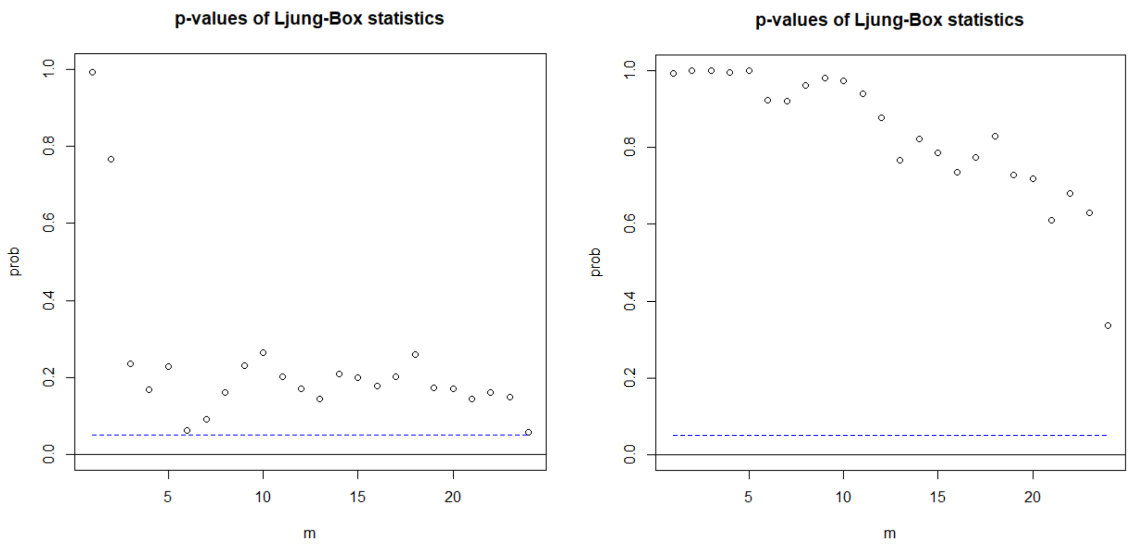

Figure 5-3: P-value Plot for the Ljung-Box Test

In Figure 5-3, from left to right are the VARMA(2,1) and VARMA(3,1) models. It can be more intuitively observed that the p-values for the Ljung-Box test of the VARMA(3,1) model are generally higher than those of the VARMA(2,1) model. If a 99% confidence level is chosen, the VARMA(2,1) model would have several points (e.g., at the fourth lag) where the null hypothesis is rejected, whereas the VARMA(3,1) model does not exhibit any such rejections.

In summary, the final model selected in this paper is VARMA(3,1). The specific parameters of the model are obtained using the maximum likelihood estimation method as follows:

$$\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} = \begin{pmatrix} 0.452 & 0.314 \\ 0.506 & 0.295 \end{pmatrix} \begin{pmatrix} z_{1,t-1} \\ z_{2,t-1} \end{pmatrix} + \begin{pmatrix} 0.0447 & 0.0875 \\ 0.1562 & -0.0868 \end{pmatrix} \begin{pmatrix} z_{1,t-2} \\ z_{2,t-2} \end{pmatrix} + \begin{pmatrix} 0.000 & 0.0741 \\ 0.234 & -0.1070 \end{pmatrix} \begin{pmatrix} z_{1,t-3} \\ z_{2,t-3} \end{pmatrix} + \begin{pmatrix} a_{1t} \\ a_{2t} \end{pmatrix} + \begin{pmatrix} 0.657 & 0.222 \\ 0.405 & 0.562 \end{pmatrix} \begin{pmatrix} a_{1,t-1} \\ a_{2,t-1} \end{pmatrix} \quad (29)$$

where the sequence $\{a_{it}\}$ is a vector that follows a normal distribution.

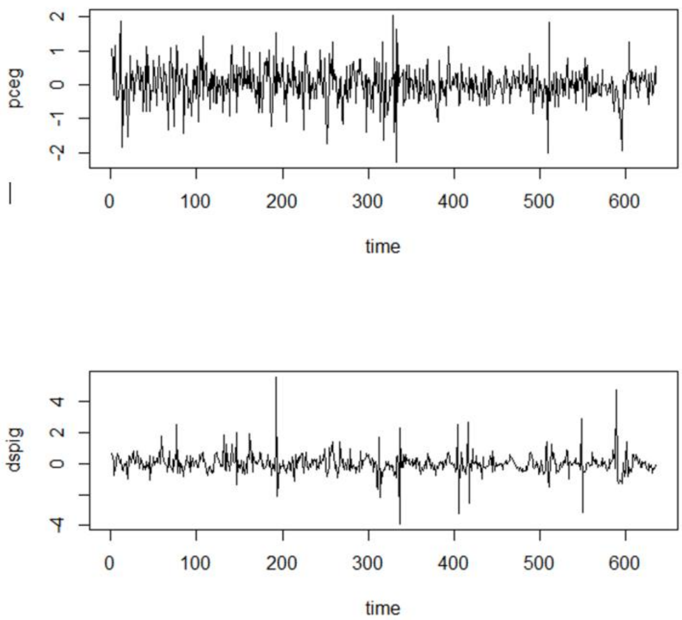

Figure 5-4: Residual Plot

Figure 5-4 indicates that the two time series variables are relatively stable over the observation period, fluctuating around a certain mean without any obvious trend or seasonality. However, there are some outliers present. This paper does not conduct a specialized analysis of these outliers. The residual series of the model appears to be white noise, suggesting that the model may have adequately fitted the data.

	F-statistic	p-value
$z2 \sim Lags(z2, 1{:}1) + Lags(z1, 1{:}1)$	5.0520	0.02494
$z2 \sim Lags(z2, 1{:}1)$	9.0171	0.00278
$z1 \sim Lags(z1, 1{:}1) + Lags(z2, 1{:}1)$	11.1422	0.0008933
$z1 \sim Lags(z1, 1{:}1)$	3.3293	0.0685267

Table 5-3: Granger Causality Test

As shown in Table 5-3, the second row indicates that the F-statistic is 5.0520 with a corresponding p-value of 0.02494, which is less than 0.05 but greater than 0.01. At the 5% significance level, we can reject the null hypothesis of Model 2, concluding that the lagged terms of z_1 significantly contribute to the prediction of z_2. The third row assesses whether the time series z_1 significantly contributes to z2 given the lagged terms of z_2. Similarly, the 4th and 5th rows can be interpreted in the same manner. Therefore, z_1 is a cause of z_2, but z_2 is not a cause of z_1.

VI. CONCLUSION

This paper provides a detailed analysis of the multivariate data composed of personal consumption expenditure and personal disposable income. Given the non-stationarity of the data, logarithmic differencing was initially applied. Rather than relying on automatically identified models, this study compares potential models and selects two superior ones. Subsequently, by comparing the autocorrelation of residuals, the optimal model is determined to be VARMA(3,1). The parameters of this model are estimated using maximum likelihood estimation. Incorporating Granger causality analysis, the study concludes that personal consumption expenditure is a cause of personal disposable income, but the reverse is not necessarily true.

REFERENCES

[1]. Yule G U. On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers[J]. Philosophical Transactions of the Royal Society of London, 1927, 226(226): 257-298.
[2]. Walker G T. Seasonal weather and its prediction[J]. Nature, 1993, 132: 805-808.
[3]. Box G E P, Jenkins G M. Time series analysis: forecasting and control[J]. Journal of the Operational Research Society, 1971, 37(2): 238-242.
[4]. Burg J P. Maximum entropy spectral analysis [J]. Society of Exploration Geophysicists,1967: 85-132.
[5]. Cryer, Jonathan D. and Kung-Sik Chan. "Time Series Analysis: With Applications in R." (2010).
[6]. Tanaka, Katsuto. "An asymptotic expansion associated with the maximum likelihood estimators in ARMA models." *Journal of the Royal Statistical Society Series B: Statistical Methodology* 46.1 (1984): 58-67.
[7]. Akaike H. Autoregressive model fitting for control [J]. Annals of the Institute of Statistical Mathematics, 1971, 23(1):163-180.
[8]. Chuang C S, Chan N H. Empirical likelihood for autoregressive model, with applications to unstable time series [J]. Statistica Sinica, 2002, 12(2):387-407.
[9]. Chan N H, Ling S. Empirical likelihood for GARCH models [J]. Econometric Theory, 2006, 22(3):403-428.
[10]. Zhao Z W, Wang D H. Empirical likelihood for an autoregressive model with explanatory variables [J]. Communications in Statistics-Theory and Methods, 2011, 40(3):559-570.
[11]. Sims C A. Macroeconomics and Reality[J]. Econometrica, 1980, 48(1): 1-48.
[12]. LÜTKEPOHL, H., AND D. S. POSKITT (1996): "Specification of Echelon-Form VARMA Models," Journal of Business and Economic Statistics, 14, 69–80.
[13]. LÜTKEPOHL, H. (1991): Introduction to Multiple Time Series Analysis. Springer-Verlag, Berlin.
[14]. Tiao G C, Box G E P. Modeling multiple time series with applications[J]. Journal of the American Statistical Association, 1981, 76(376): 802-816.
[15]. Engle R F, Granger C W J. Co-integration and error correction: representation, estimation, and testing[J]. Econometrica, 1987, 55(2): 251-276.
[16]. Hendry D F, Anderson G J. Testing dynamic specification in small simultaneous systems: an application to a model of Building Society behaviour in the United Kingdom[J]. Cowles Foundation Discussion Papers, 1977, 57(11): 89-95.
[17]. BOSQ, D. (1998): Nonparametric Statistics for Stochastic Processes - Estimation and Prediction, no. 110 in Lecture Notes in statistics. Springer-Verlag, Berlin, second edn.
[18]. DOUKHAN, P. (1995): Mixing - Properties and Examples, no. 85 in Lecture Notes in Statistics. Springer-Verlag.
[19]. P. Whittle, Thesis : Hypothesis Testing in Time Series Analysis, Almquist and Wiksell, Upsala, 1951.
[20]. J. Durbin, Efficient estimation of parameters in moving-average models, Biometrika 46 (1959) 306–316.
[21]. E.J. Godolphin, A direct representation for the large sample maximum likelihood estimator of a Gaussian autoregressive moving average process, Biometrika 71 (1984) 281–289.
[22]. E.J. Hannan, J. Rissanen, Recursive estimation of mixed autoregressive-moving average order, Biometrika 69 (1982) 81–94.
[23]. S. Koreisha, T. Pukkila, Fast linear estimation methods for vector moving average models, J. Time Series Anal. 10 (1989) 325–339.
[24]. G. Tunnicliffe-Wilson, The estimation of parameters in multivariate time series models, J. R. Stat. Soc Ser. B 35 (1973) 76–85.
[25]. G.C. Reinsel, S. Basu, S.F. Yap, Maximum likelihood estimators in the multivariate autoregressive moving average model from a generalized least squares viewpoint, J. Time Series Anal. 13 (1992) 133–145.
[26]. R.F. DeFrutos, G.R. Serrano, A generalized least squares estimation method for invertible vector moving average models, Econom. Lett. 57(1997) 149–156.
[27]. Tsay, Ruey S.. "Multivariate Time Series Analysis: With R and Financial Applications." (2013).
[28]. Dunsmuir, W. and Hannan, E. J. (1976). Vector linear time series models. Advanced Applied Probability, 8: 339- 364.
[29]. Hannan, E. J. and Deistler, M. (1988). The Statistical Theory of Linear Systems. John Wiley & Sons, Inc, New York. Republished in 2012 as Classics in Applied Mathematics, Society for Industrial and Applied Mathematics.