The Scientific Journal of Business and Finance

https://caf.journals.ekb.eg

Assessing the Dynamics of Interaction in Multiple Time-Varying Variables Via the Symmetric Lag Vector AutoRegressive (VAR) Model

Sohair F, Higazi ,[a] Hani A, Khedr[b] Sarah F, Aboud [c]

[a] Professor of Applied Statistics, Faculty of Commerce, Tanta University
[b] Lecturer, Statistics Department, Faculty of Commerce, Damanhour University
[c] Lecturer, Statistics Department, High Institute of Computers, Information and Management Technology, Tanta.

Published online: **December 2024.**
To cite this article: Higazi, Sohair F, Khedr, Hani A, & Aboud, Sarah F. **Assessing the Dynamics of Interaction in Multiple Time-Varying Variables Via the Symmetric Lag Vector AutoRegressive (VAR) Model** Journal of Business and Finance, 44, (4),162-181

DOI: 10.21608/caf.2024.392536

*Corresponding author: sarahfathey2013@gmail.com

Assessing the Dynamics of Interaction in Multiple Time-Varying Variables Via the Symmetric Lag Vector AutoRegressive (VAR) Model

Sohair F Higazi
Professor of Applied Statistics, Faculty of Commerce, Tanta University
Hani A, Khedr
Lecturer, Statistics Department, Faculty of Commerce, Damanhour University
Sarah F Aboud
Lecturer, Statistics Department, High Institute of Computers, Information and Management Technology, Tanta.

Article History
Received 1 october2024, Accepted 28 october 2024, Available online December 2024.

Abstract
This paper aims to investigate the effects and interactions of time series data when using the bivariate and the multivariate symmetric VAR model; and also, to evaluate the impact of relationship on shock and on short and long-term Impulse response analysis. Data used in the analysis were World Bank data for Egypt for the years from 1990 to 2020. Two symmetric VAR models were used to analyze the relationship between inflation (INF) and some key macroeconomic indicators that include the trade balance deficit (TBD), gross domestic product (GDP), government expenditure (GEX), and foreign investments (FV). The first model used was the bivariate VAR model, which focuses on inflation (INF) and trade balance deficit (TBD), and then, applying a multivariate VAR model containing all other variables. The primary objective was to assess the accuracy and explanatory power of these models in forecasting inflation based on Akaike information criterion (AIC), Schwarz criterion (SC), and also on Hannan-Quinn information criterion (HQ).
Statistical results show that the multivariate VAR model significantly improves forecasting accuracy; the bivariate model achieved lower values for AIC and SC indicating a simpler structure, but the multivariate model demonstrates more robust performance, particularly in capturing the long-term effects of additional variables. Both models revealed that inflation is primarily driven by its own shocks, with only a minor contribution from the trade balance deficit. However, the multivariate model provided a broader explanation of inflation's variance, as financial variables became increasingly influential over time. Impulse response analysis indicated that both models exhibited similar short-term reactions to shocks, though the effects diminished faster in the multivariate model. The study concludes that the bivariate VAR model is more appropriate for analyzing the effects of a variable on itself and short-term shock impacts, while the multivariate VAR model is more suited for variance decomposition and understanding complex economic dynamics. The recommendations emphasize the importance of advanced statistical models like the multivariate VAR to enhance forecasting accuracy and long-term Impulse

Response Analysis economic decision-making. And the univariate VAR model for short-term Impulse Response analysis.

Key Words: Autoregressive Models; Variance Decomposition; Box-Jenkins Models; Asymmetric Lags Inflation; Economic Growth; Impulse Shock response; Granger Causality; Optimal Lag.

1. Introduction

The VAR model relates variables that cannot be differentiated from being internal or external variables, and they are all considered internal variables. The VAR model differs from univariate autoregressive (AR) models because the AR model deals with only one variable and it measures how the variable at a specified point in time is related to its previous values; while the VAR model deals with at least two variables; allows feedback between the variables, and thus allows for the analysis of multiple time series variables together. VAR model describes the joint behavior of multiple time series variables based on their lagged values, and thus, it is considered as a multivariate extension of autoregression (AR) models. The VAR models are either "symmetric" in lags, meaning that each time series has the same lag length, or asymmetric, meaning that each time series variable has different lag structure. VAR assumes that each variable linearly depends on past values of itself and other variables in the system. The key assumptions of a VAR model are stationarity, linearity, and a constant covariance matrix of the error terms. Additionally, VAR models assume that variables in the system have a contemporaneous effect on each other, capturing the dynamic interactions within the system. Thus, the VAR model is a systematic but flexible approach for capturing complex real-world behavior that improves forecasting, and it captures the simultaneous dynamics of time series data and help in predicting the likelihood of occurrence of a time series of stock processes, weather forecasts, traffic conditions, and all other events affected by historic occurrences.

The VAR model differs than Granger causality (Granger, 1969); the VAR model is used to investigate relationships between the variables, while Granger causality examines the relationship and their direction between time series. Granger causality tests if the prediction of one time series is improved by a second time series. However, the VAR model consists of a set of equations, where each equation represents one variable as a function of its own lagged values and the lagged values of other variables in the system.

VAR models are used in econometrics for modelling correlated variables with bi-directional relationships, however Stock and Watson (2001) analyze systems with multiple interrelated variables. VAR models in economics were made popular by Sims (1980). The definitive technical reference for VAR models is Lutkepohl (1991), and updated surveys of VAR techniques are given in Watson (1994) and Lutkepohl (1999) and Waggoner and Zha (1998).

Several applications for modelling correlated time series using VAR models in the various disciplines are found in the literature, to name a few: Longmore (2020), Bayraci et al (2011), Rodolfo Marcano (2021), Eliezer Bose et al (2017) and Nalita et al (2021).

VAR modeling using Bayesian approach is found in Musyoki et al, (2023), Feihan et al (2018). VAR model with machine learning algorithms is discussed in Huang et al (2021), Li and Yuan (, 2022, 2024), Agusti and Costa (2021); Aydin and Cavdar (2015); Nichlson et al (2022), and a comparison between machine learning VAR and classical Var is given in (Aguilar, undated).

The paper aims to investigate the effects and interactions of time series data when using the bivariate and the multivariate symmetric VAR model; and also, to evaluate the impact of relationship on shock and on short and long-term Impulse response analysis. Data used in the analysis were World Bank data for Egypt for the years from 1990 to 2020.

Several studies investigated inflation; each study dealt with different variables and with different statistical models. Most of the econometric studies on inflation use ARIMA and or regression models using OLS; the study (Kheir-El-Din, 2008) used temporal data during two periods, the first before 1990/1991, while the second period comes after 1990/1991, where the growth equation with inflation was estimated by cross-sectional regression and use Ordinary least square to estimate regression equation. Rihan (2018) investigated the impact of inflation on growth rate use simultaneous equations models; Hammam (2010) uses annual time series analysis to assess the significance of several important variables on economic growth in Egypt; ARIMA model was applied to a non-stationary time series and used OLS method for parameter estimation. Monem (2018) analyzed Egypt's inflation dynamics from 1980 to 2009, comparing them to global trends, the VAR model along with some other econometric methods across three sub-periods were used and concluded that Egypt's inflation more persistent and responsive to supply shocks. Abd El-Aal (2023) studied Egypt's inflation via various machine learning algorithms, the study showed that the Gradient Boosting (G.B.) algorithms is the most accurate, and that there are positive relationships between inflation and several factors including government expenditure and GDP growth, and a negative relationship with household expenditure and the external trade balance. Kamal (2023) examined Egypt's economic growth from 1991 to 2019 using the augmented Solow model and quantile regression; the study showed that productivity has a significant impact on higher stages of economic growth, while capital and human capital contributions decrease over higher income segments, and that the weak growth is attributed to low savings rates and high population growth, highlighted the importance of human capital, productivity, and capital influenced by these factors. Helmy (2008) examined the short-run and long-run relationships between Egypt's budget deficit, its financing sources, and inflation from 1981/82 to 2005/06, the VECM analysis indicates significant two-way dynamic interaction among the budget deficit, government credit, exchange rate, and inflation. Nabil (2022) used The AI Long Short-Term Memory (LSTM) models to forecast Egypt's GDP in a univariate (GDP) and a multivariate (GDP along with, unemployment, inflation rate) time series, with integrated charts for group policy selection, the study concluded that multivariate models with GDP and inflation rate outperformed both univariate models and those with three indicators in terms of RMSE and R-squared.

The reviewed studies highlight the impact of inflation, monetary policy, and other economic factors on growth and public spending using various econometric models such as VAR, VECM, and ARIMA. However, there is a gap in comparing the effectiveness of

two-variable versus multivariate VAR models in estimating inflation dynamics, justifying the need for further research to make this comparison on statistical grounds.
The current study is applied to World Bank data for Egypt, in the time period from 1990 to 2020 (the latest release dated September 25, 2023), variables used in the study are on inflation, gross domestic product, government spending, foreign investment, exports, and imports. One of the reasons for choosing the field of application is that the relationship between inflation and trade balance deficit is considered an important and controversial issue. In this research, the relationship between Inflation (INF) and Trade Balance Deficit (TBD) is studied in the Vector Autoregressive (VAR) model, which represents the first model, and compared to the VAR model, which contains many variables, the effect of shocks for each variable on the study variable (inflation) is studied in the two models and variance decomposition of INF in two Models and Study of the impact of shocks.

The study has five objectives: 1) to give an overview of the VAR models; 2) to apply the VAR model to a two variables system (inflation (INF) and Trade Balance Deficit (TBD)); 3) to evaluate the VAR model by introducing more variables to the model in (2) above; 4) to compare the fitness of the models in (2) and (3) above; and 5) to study the effect of inflation shocks and impulse response on the other variables.

The paper is divided into five sections (other than the introduction); Section 2 introduces the bivariate and multivariate VAR systems; Section 3 covers estimation techniques and forecasting of the VAR model; Section 4 gives results of VAR Model when applied to real data; and in Section 5 conclusions and recommendations are given.

2. The Vector Autoregressive (VAR) Model

The VAR model is a multi-step process model that involves a) Specification and estimation; b) checking and revising the model; c) Forecasting, and d) Structural analysis. There are three broad types of VAR models, the reduced, the recursive, and the structural form. The simplest model is the reduced VAR models when each variable is considered as a function of its own past values, and the past values of other variables in the model. The drawback of the reduced model is that the associated variables are not related to one another; and the error terms are correlated across equations (Eric, 2021).
To capture the dynamic interactions within the VAR model, it assumes constant covariance matrix of the error terms and contemporaneous effect of each variable in the system. Thus, in addition to the assumptions of stationarity and linearity, the VAR Model assumes:

1. $E(e_t) = 0$, i.e, every term has a zero mean.
2. $E(e_t, e`_t) = \Omega$.
   $\text{where } \Omega \text{ is the contemporanous covariance matrix of the k} \times \text{k positive semi} - \text{matrix}$
3. $E(e_t, e`_{t-k}) = 0$ for any non-zero error term, i.e., no serial correlation.

2.1 Model Specification

When specifying the VAR model, we need to specify the number of endogenous variables and the number of the autoregressive terms in the model; the order of the VAR model depends on the latter. Thus, for two endogenous variables and two autoregressive terms is referred to as Bivariate VAR (2). And for 3 endogenous variables with 4 autoregressive terms is referred to as multivariate.

The ARMA(p) model with t=1,2,…n; and lag(p) model takes the following form:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots .. + \phi_p y_{t-p} + \varepsilon_t \qquad (1)$$

In matrix form, equation (1) could be expressed as:

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & y_0 & \cdots & y_{1-p} \\ 1 & y_1 & \cdots & y_{2-p} \\ . & . & ..., & \cdots \\ . & . & .., & \cdots \\ 1 & y_{n-1} & \cdots & y_{n-p} \end{bmatrix} \begin{bmatrix} \phi_0 \\ \phi_1 \\ \cdots \\ \cdots \\ \phi_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdots \\ \cdots \\ \varepsilon_n \end{bmatrix} \qquad (2)$$

Where

y_t :data of Y in time t, t=1,2, …, n;

y_{t-p} : data of Y in time (t-p) period.,

ε_t : error in time t; ; ϕ_0 $a\ constant\ for\ the\ mean$;

ϕ_i : $autoregressive\ coefficient\ \ i = 0, 1,2, \dots p$

In the bivariate case, equation (1) with p=1, becomes a VAR(1) model :

$$y_{1,t} = \phi_{10} + \phi_{11} y_{1,t-1} + \phi_{12} y_{2,t-1} + \varepsilon_{1,t}$$

$$y_{2,t} = \phi_{20} + \phi_{21} y_{1,t-1} + \phi_{22} y_{2,t-1} + \varepsilon_{2,t} \qquad (3)$$

Equation (3) is with an intercept $\phi_{k,0}$; $k = 2$ is the number of variables in the system, and symmetric lag (p=1 for both variables); shown in matrix format for VAR(1) as:

$$\boldsymbol{Y}_{n\times 2} = \boldsymbol{W}_{n\times 3}\, \boldsymbol{\phi}_{3\times 2} + \boldsymbol{E}_{n\times 2} \qquad (4)$$

Where,

$$Y_{n\times 2}=\begin{bmatrix} Y_{1,1} & Y_{2,1} \\ Y_{1,2} & Y_{2,2} \\ \cdots & \cdots \\ Y_{1,n} & Y_{2,n} \end{bmatrix} \quad W_{n\times 3}=\begin{bmatrix} 1 & y_{1,1-1} & y_{2,1-1} \\ 1 & y_{1,2-1} & y_{2,2-1} \\ . & . & \cdots \\ . & . & \cdots \\ 1 & y_{1,n-1} & y_{2,n-1} \end{bmatrix} \quad \phi_{3\times 2}=\begin{bmatrix} \phi_{10} & \phi_{20} \\ \phi_{11} & \phi_{21} \\ \phi_{12} & \phi_{22} \end{bmatrix}$$

$$E_{n\times 2}=\begin{bmatrix} \varepsilon_{1,1} & \varepsilon_{2,1} \\ \varepsilon_{1,2} & \varepsilon_{2,2} \\ \cdots & \cdots \\ \varepsilon_{1,n} & \varepsilon_{2,n} \end{bmatrix}$$

Equation (4) is extended to the multivariable case for VAR(1) as follows :

$$\boldsymbol{Y}_{n\times k}=\boldsymbol{W}_{n\times(k+1)}\boldsymbol{\phi}_{(k+1)\times k}+\boldsymbol{E}_{n\times k} \tag{5}$$

Where,

$$Y_{n\times k}=\begin{bmatrix} Y_{1,1} & \cdots & Y_{k,1} \\ Y_{1,2} & \cdots & Y_{k,2} \\ \cdots & \cdots & \cdots \\ Y_{1,n} & \cdots & Y_{k,n} \end{bmatrix} \quad W_{n\times(k+1)}=\begin{bmatrix} 1 & y_{1,1-1} & \cdots & y_{k,1-1} \\ 1 & y_{1,2-1} & \cdots & y_{k,2-1} \\ . & . & \cdots & \cdots \\ . & . & \cdots & \cdots \\ 1 & y_{1,n-1} & .. & y_{k.n-1} \end{bmatrix}$$

$$\phi_{(k+1)\times k}=\begin{bmatrix} \phi_{10} & \cdots & \phi_{k0} \\ \phi_{11} & \cdots & \phi_{k1} \\ \cdots. & \cdots & \cdots. \\ \phi_{1k} & .. & \phi_{kk} \end{bmatrix} \quad E_{n\times k}=\begin{bmatrix} \varepsilon_{1,1} & \cdots & \varepsilon_{k,1} \\ \varepsilon_{1,2} & \cdots. & \varepsilon_{k,2} \\ \cdots & \cdots. & \cdots \\ \varepsilon_{1,n} & \cdots. & \varepsilon_{k,n} \end{bmatrix}$$

k: number of variables; n= number of time periods. The matrix form in (4) and (5) could be extended to VAR(p) for $p < n$ where:

$$\boldsymbol{Y}_{\boldsymbol{n}\times\boldsymbol{k}}=\boldsymbol{W}_{\boldsymbol{n}\times(\mathbf{1}+\boldsymbol{pk})}\boldsymbol{\phi}_{(\mathbf{1}+\boldsymbol{pk})\times\boldsymbol{k}}+\mathbf{E}_{\boldsymbol{n}\times\boldsymbol{k}} \tag{6}$$

2.2 Optimal Lag structure (p)

Determination of the lag structure for the VAR model is very essential in the model specification (Braun and Mittnik,1993). Overfitting causes an increase of mean-square forecasts and underfitting the lag length often generates autocorrelated errors (Lütkepohl ,1993); and the accuracy of forecasts varies substantially for alternative lag length (Hafer and Sheehan (1989) . Symmetric lag VAR models (same lag for all variables) could be estimated using OLS estimation procedure, since the specification of all equations of the model is the same (Ozcicek and McMillin, Louisiana State University working paper).

Several criteria could be used to determine the optimal lags length. The analysis often starts with lag 0 (only constants), then 1, 2, 3, …, lags; and comparing one or more of the following criteria, at each lag length, the s the criterion the better the lag length; and thus, reaching and choosing the optimal lag length. One way to decide on lag length is the use of Wald test statistics (Nalita et al. 2021; Davidson and Mackinnon, 1993) to test whether all the coefficients at each lag is zero. Once the lag structure has been obtained and has serially uncorrelated errors, one can generate forecasts. The lag length is determined under five criteria (Zivot , 2006):

i)The final prediction error (FPE) criteria.

$$FPE(p) = \left[\frac{n+kp+}{n-kp-}\right]^2 \times |\Sigma_{\hat{u}\hat{u}}(p)\ | \qquad (7)$$

ii)The Akaike Information criterion (AIC):

$$AIC\ (p) = ln|\Sigma_{\hat{u}\hat{u}}(p)\ | + \ (k + pk^2)\frac{2}{n} \qquad (8)$$

iii) The Schwartz-Bayesian Information Criteria (SBIC or SC):

$$SC(p) = ln|\Sigma_{\hat{u}\hat{u}}(p)\ | + (k + pk^2)\frac{\ln(n)}{n} \qquad (9)$$

iv) The Hannan-Quinn Criterion (HQ):

$$HQ(p) = ln|\Sigma_{\hat{u}\hat{u}}(p)\ | + (k + pk^2)\frac{2\ln\ln(n)}{n} \qquad (10)$$

v) The likelihood Ratio test (LR)

The LR test compares the goodness of fit between two models, one with fewer lags than the other, computed as:

$$\text{LR} = -2(\text{Log maximized likehood with few lags} - \text{Log maximized likehood with more lags}) \qquad (11)$$

In the first four criteria, n is the length of the time series; k = number of variables, p = number of lags, and

$|\Sigma_{\hat{u}\hat{u}}(p)|$ is the determinant of the variance covariance matrix of the estimated residuals (Nalita et al. 2021).

3. Estimation and Forecasting

The common methods for parameter estimation are Ordinary Least Square (OLS) and Maximum Likelihood Estimation (MLE). Estimation by ordinary least squares yields efficient parameter estimates (Nalita et al 2021). Estimates using OLS is obtained per equation and a one-period-ahead forecast is computed to obtain the two-period-ahead forecasts; continue by iterating to obtain forecasts of all variables in the VAR farther into the future. While OLS minimizes the square error function, MLE maximizes the log-likelihood function. The required function for both OLS and MLE method that provide the same results (using equations (4) (5) or (6)) is (Nalita et al 2021):

$$\widehat{\boldsymbol{\phi}} = (\boldsymbol{W^T W})^{-1}\boldsymbol{W^T Y} \qquad (12)$$

However, while the OLS requires that: a) variables in the system are stationary; 2) errors have a mean of zero; 3) no perfect collinearity; and 4) no outliers. Under these assumptions, the OLS estimates are consistent, and efficient when the disturbances have mean zero, constant variance, the MLE requires (Wei, 2006) that the error term $\varepsilon_i \sim N(0, \sigma^2)$. The number of parameters to be estimated in the VAR model is $(1 + pk) \times k = k + pk^2$. When k=5 and p=2, the number of estimated coefficients is 55, and the more the coefficient, the higher the error of prediction. Thus, it is recommended to keep k and p as small as possible.

Parameter Significance for the Var(p) model, hypothesis is evaluated using Wald t-statistic and p-values to test the significance of the estimator, the null and the alternative hypotheses are: H_0: $\phi_i = 0$ vs H_1: $\phi_i \neq 0$ where the Wald t- test is $t = \frac{\widehat{\phi}_i}{SE(\widehat{\phi}_i)}$ (Nalita et al. 2021, Davidson and Mackinnon, 1993).

Forecasts are generated for VAR models using an iterative forecasting algorithm; The VAR function is included in many statistical data analysis packages, to name a few: EViews, R (VAR), Python (VARs); Stata (Var); SAS (VARMAX), EViews (VAR); MATLAB (varm); Regression analysis of time series (SYSTEM). Using OLS, the following is performed:

1. Estimate the VAR model using OLS for each equation.
2. Compute the one-period-ahead forecast for all variables.
3. Compute the two-period-ahead forecasts, using the one-period-ahead forecast.
4. Iterate until the h-step ahead forecasts are computed.

Also, VAR routines in all statistical packages are useful to test whether one variable is useful in forecasting another variable or not; it gives impulse response analysis where the response of one variable witnesses a sudden but temporary change in another variable; and thus, assessing the dynamic interactions and transmission mechanisms within the system. Forecasts error variance decomposition is also displayed in the output of all packages, where the proportion of the forecast variance of each variable is attributed to the effects of the other variables.

4. Application

The bivariate and multivariate VAR model is applied to World Bank data for Egypt, for the period from 1990 to 2020; variable used are: trade transactions which is exports (EXP) and imports (IMP), per capita GDP, government expenditure (GEX) expressed as a percentage of GDP, Foreign direct investments in billion \$ (FV), Inflation (INF) expressed as the annual average according to consumer prices, balance of trade is the difference between exports and imports over for all time periods.

The VAR system is used to clarify the relationship between these variables, to find out the extent of the effect of increasing the variables in the autoregression model, and to study the effect of shocks and variance decomposition.

Two models are applied to these data, the bivariate model to reach the relationship between Inflation (INF) and Trade Balance Deficit (TBD); and a multivariate model (VAR) model, which contains all variables used in the study, where the effect of shocks for each variable on inflation) is studied.

Figure 1 displays how Trade Balance Deficit (TBD) affects inflation and Figure 2 shows the dynamic nature of economic variables with inflation (INF) being the most volatile, possibly influenced by changes in TBD and other economic conditions.

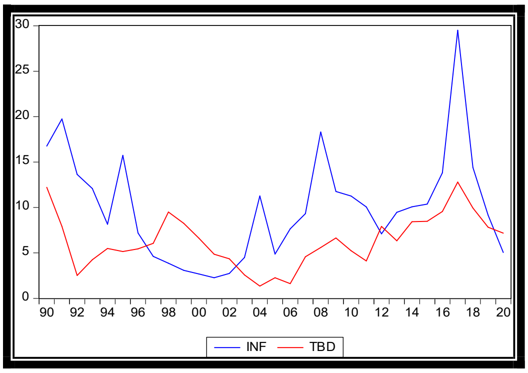

Fig.1: Volatility of Inflation and Trade Balance deficit

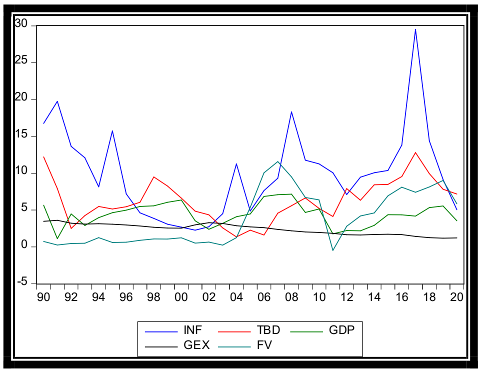

Fig:2. Volatility of all variables

Stationarity of data implies some statistical properties of the series, such as mean and variance, remain constant over time. A unit root test results are shown in Table(1) for a bivariate case (INF, TBD), where it is evident that all tests are significant (P<.05).

Table 1: Unit root Tests for Two Variables (INF and TBD)

Group unit root test: Summary
Series: INF, TBD
Sample: 1990 2020

Method	Statistic	Prob.**	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu t*	-1.76080	0.0391	2	60
Null: Unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat	-2.20263	0.0138	2	60
ADF - Fisher Chi-square	11.5325	0.0212	2	60
PP - Fisher Chi-square	11.8393	0.0186	2	60

The unit root tests were conducted, also, on the five variables: INF, DTB, GDP, GEX, and FV, over the period from 1990 to 2020. Table (2) gives stationary Unit root test results, for all variable series; it is evident that all tests are significant ($p< .05$), and thus the VAR model can be estimated.

Table 2: Unit root Tests for Multiple Variables

Group unit root test: Summary
Series: INF, DTB, GDP, GEX, FV
Sample: 1990 2020

Method	Statistic	Prob.**	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu t*	-1.01675	0.04933	5	148
Null: Unit root (assumes individual unit root process)				
Im, Pesaran and Shin -Wstat	-2.07086	0.0381	5	148
ADF - Fisher Chi-square	20.6585	0.0236	5	148
PP - Fisher Chi-square	21.5592	0.0175	5	150

** Probabilities for Fisher tests are computed using an asymptotic Chi -square distribution. All other tests assume asymptotic normality.

4.1 Optimal Lags for the Bivariate and Multivariate VAR

To determine the optimal number of lags for each of the VAR models, Equations [7, 8, 9, 10, 11)] are computed. Table (3) gives the optimal lag for the bivariate VAR model (Inflation and TBD); and Table (4) gives the optimal lag for the Multivariate VAR model, where all variables under study are used.

Table 3: Optimal Lag Length for a Bivariate VAR Model (INF and TBD)

VAR Lag Order Selection Criteria
Endogenous variables: INF, TBD
Exogenous variables: C
Sample: 1990 2020
Included observations: 29

Lag	Log L	LR	FPE	AIC	SC	HQ
0	-158.9841	NA	227.3865	11.10235	11.19665	11.13188
1	-144.9314	25.19800*	113.8396*	10.40906*	10.69195*	10.49766*
2	-143.8062	1.862405	139.5642	10.60732	11.07880	10.75498

Tables (3) and (4) show that Lag 1 is the optimal lag length according to the LR, FPE, AIC, SC, and HQ criteria. It provides the lowest values across these measures, suggesting that it balances model complexity and fits most effectively. This lag length offers a good compromise between predictive accuracy and avoiding overfitting, ensuring a robust model.

Table 4: Optimal Lag Length for Multiple Variables VAR Model

VAR Lag Order Selection Criteria
Endogenous variables: INF TBD GEX GDP FV
Exogenous variables: C
Sample: 1990 2020
Included observations: 29

Lag	Log L	LR	FPE	AIC	SC	HQ
0	-290.8399	NA	499.4470	20.40275	20.63849	20.47658
1	-194.2803	153.1634*	3.700724*	15.46761*	16.88205*	15.91060*
2	-176.1545	22.50103	7.040933	15.94169	18.53484	16.75383

4.2 Estimation of the VAR Models

Using OLS estimation technique (Equation (12)) and Equation (4), where: INF_t represent $Y_{1.t}$, TBD_t represent $Y_{2.t}$ & lag = 1; equation (3) is estimated as:

$$\begin{bmatrix} \widehat{INf} \\ \widehat{TBD} \end{bmatrix} = \begin{bmatrix} 3.845 \\ 2.06 \end{bmatrix} + \begin{bmatrix} .4760 & .7441 \\ -.0052 & .6521 \end{bmatrix} \begin{bmatrix} INF_{-1} \\ TDB_{-1} \end{bmatrix}$$

And $R^2_{INF} = .2920 \quad R^2_{TBD} = .4920$

Thus, inflation of current period is highly affected by the trade balance deficit of a previous period, and trade balance deficit in a current period is also highly affected by its previous period; and inflation and TBD of a previous period contribute 49.20% in the variability of TBR in the current period.

Using OLS estimation technique (Equation[12]) and Equation [5] , where: where, INF_t represent $Y_{1.t}$, TBD_t represent $Y_{2.t}$, GDP_t represent $Y_{3.t}$, GEX_t represent $Y_{4.t}$, FV_t represent $Y_{5.t}$ & lag=p = 1. We get the following estimation equations:

$$\begin{bmatrix} \widehat{INF_t} \\ \widehat{TBD_t} \\ \widehat{GDP_t} \\ \widehat{GEX_t} \\ \widehat{FV_t} \end{bmatrix} = \begin{bmatrix} -7.27 \\ 6.96 \\ 1.21 \\ .06 \\ 4.39 \end{bmatrix} + \begin{bmatrix} .31 & .79 & -1.48 & 4.83 & 1.37 \\ .02 & .46 & .32 & -1.95 & -.16 \\ -.01 & -.01 & .22 & .72 & .22 \\ -.01 & .01 & .03 & .92 & -.02 \\ .01 & -.16 & -.36 & -.46 & .86 \end{bmatrix} \begin{bmatrix} INF_{t-1} \\ TBD_{t-1} \\ GDP_{t-1} \\ GEX_{t-1} \\ FV_{t-1} \end{bmatrix}$$

The $R^2_{INF} = .5039$; $R^2_{TBD} = .6271$; $R^2_{GDP} = .3028$

$R^2_{GEX} = .9634$ *and* $R^2_{FV} = .7273$. It is clear that inflation in a current period is highly affected by the government spending (GEX) in a previous period and by foreign investments (FV);trade balance deficit (TBD) is negatively correlated with government spending and foreign investments; GEX in a current period is affected the most by the spending in a previous period, and its R^2 value indicates that the contribution of INF, TBD, GDP, GEX, and FV contribute 96.34% of its variability; all R^2 values indicate that interaction exists between all variables, however, R^2 value for GDP is a bit low.

4.3 Fitness of the VAR Models

The bivariate and the multivariate VAR models were compared. Table (5) gives fitness criteria for the two models.

Table 5: VAR Models Fitness

Criterion	Bivariate		Multivariate				
	INF	BTD	INF	BTD	GDP	GEX	FV
R^2	.29	.49	.5039	.6271	.3028	.9634	.7273
F statistic	5.57	12.98	4.87	8.07	2.08	126.35	12.80
LL	-90.93	-61.62	-84.98	-56.94	-50.33	17.54	-61.64
AIC	6.22	4.31	6.06	4.19	3.76	-.76	4.52
SC	6.36	4.48	6.34	4.48	4.04	-4.89	4.79

The p-value for all models is less than .05, and thus all models are significant; however, GEX is highly significant, means that it is affected very highly with other variables (used in the study). The AIC and SC criterion for the variable "INFlation", are less under the multivariate VAR models than the bivariate VAR model. Thus, it is concluded that increasing the number of variables in the VAR model increased the explanatory power of the model.

4.4 Variance Decomposition for the variable Inflation

The variance decomposition table provides insights into the proportion of the forecast error variance of inflation (INF) that can be attributed to its own shocks and to shocks in other variables over different periods. The focus is on how the inclusion of additional variables changes the explanatory power of INF itself and on other variables. The variance decomposition (Tables 6 and 7) reflects the contribution of each variable to the forecast error variance of inflation (INF) over different periods in a VAR model.

Table 6: Variance Decomposition for INFLATION under the bivariate model

Variance Decomposition of INF:			
Period	S.E.	INF	TBD
1	5.178010	100.0000	0.000000
2	5.804496	99.40882	0.591176
3	5.969962	98.72859	1.271408
4	6.021966	98.24139	1.758606
5	6.040552	97.95774	2.042255
6	6.047805	97.81044	2.189563
7	6.050780	97.73908	2.260925
8	6.052030	97.70603	2.293975
9	6.052560	97.69118	2.308824
10	6.052784	97.68464	2.315357

Examination of Table 6 shows that: a) INF explains 100% of its variance in period one; b) INF's self-contribution slightly decreases from 99.41% to 97.68% for the Periods 2-10; c) TBD's contribution gradually increases from 0.59% to 2.32% for the periods 1 to 10; d) the overall TBD has a minor but gradually increasing influence on INF, and finally e) the variance in INF remains mostly explained by its own shocks.

Table 7: Variance Decomposition for INFLATION under the Multivariate VAR model

Variance Decomposition of INF:						
Period	S.E.	INF	TBD	GDP	GEX	FV
1	4.597203	100.0000	0.000000	0.000000	0.000000	0.000000
2	5.799301	75.01527	2.751420	1.726534	5.872842	14.63393
3	6.143877	68.58817	3.265033	1.573342	7.237339	19.33612
4	6.264980	66.22003	3.254815	1.538396	7.607899	21.37886
5	6.320162	65.11447	3.199213	1.543173	7.712515	22.43062
6	6.350540	64.50734	3.180460	1.542272	7.738150	23.03178
7	6.369228	64.13849	3.186168	1.536978	7.739107	23.39926
8	6.381432	63.90185	3.199211	1.531737	7.732678	23.63453
9	6.389681	63.74533	3.211514	1.527826	7.724705	23.79063
10	6.395403	63.63942	3.220885	1.525095	7.717159	23.89744

Examination of Table (7) shows that:

a) INF explains 100% of its variance, for period 1.

b) The contribution of INF itself decreases more significantly for the periods from 2 to 10 compared to the two-variable model, from 75.02% to 63.64% and other variables like TBD, GDP, GEX, and FV start to play a more noticeable role:

1)TBD: Contribution starts at 2.75% and slightly increases to 3.22%.
2) GDP: Has a minimal impact (starting at 1.73% and ending at 1.53%).
3) GEX: Shows a relatively stable influence (from 5.87% to 7.72%).
4) FV: Its contribution grows from 14.63% to 23.90%, indicating a significant role in explaining INF variance
5) Overall: While INF's self-explanatory power diminishes in the presence of more variables, additional new variables take on a more substantial explanatory role. TBD still has a minor impact, similar to the two-variable model

4.5 Impulse Shock Response

The impulse response function shows the responses of Inflation (INF) and Trade Balance Deficit (TBD) to a one standard deviation shock in each variable over 10 periods, based on a Vector Autoregression (VAR) model. Figure (3) displays the impulse responses, and ±standard error bands, which provide a measure of statistical significance.

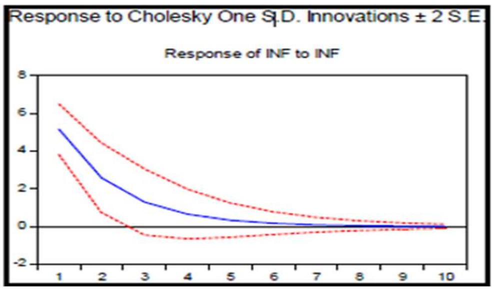

Fig. 3: Impulse Response of Inflation (Bivariate VAR)

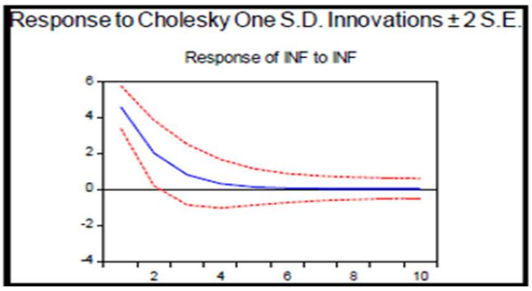

Fig. 4: Impulse response Inflation (Multivariate)

A comparison between Shock response of INF to INF for the Bivariate and the multivariate VAR Models is given in Table 8, while Table 9, gives the Shock Response of INF to TBD for the Bivariate and the multivariate VAR Models.

Table 8: Compare Impulse Shock Response of INF to INF

Compare Aspect	Bivariate VAR Model (Fig. 3)	Multiple VAR Model (Fig. 4)
Initial Impact (Short-term Response)	Strong positive impact, peaks in the first period. √	The same. √
Statistical Significance (Short-term)	Significant in the first few periods. √	The same. √
Long-term Response	Gradually declines, becomes insignificant after 7 periods.	Gradually declines, becomes insignificant after 6 periods.
General Trend	The impact diminishes over time, stabilizing at zero. √	The same √

Table 8 (Figures 3 and 4) highlights the key similarities and differences in the responses of inflation to its own shocks across the two models. Both models show a similar short-term positive response and long-term decline, but the multiple VAR model stabilizes slightly earlier. Thus, it is concluded that the bivariate VAR shows a persistent yet diminishing effect of inflation shocks on itself in the long run, and the multivariate VAR shows similar diminishing effect but stabilizes slightly earlier (by the 6th period).
Table 9 (Figures 5 and 6) shows that the bivariate model exhibits a stronger short-term response to TBD shocks, while the multiple models demonstrate a weaker effect that persists for a slightly longer period before becoming insignificant. It is concluded that the bivariate model exhibits strong but temporary effects of trade balance deficit (TBD) shocks on inflation (INF), which quickly dissolves; this effect is weaker in the multiple models but lasts slightly longer before fading.

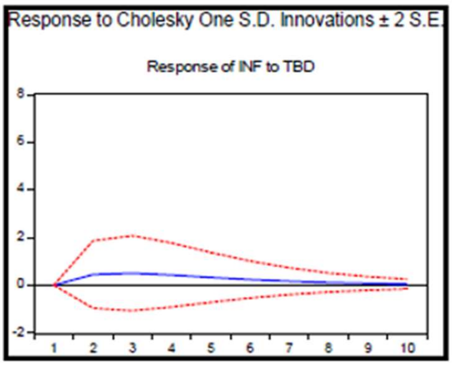

Fig. 5: Inflation Shock(Bivariate VAR)

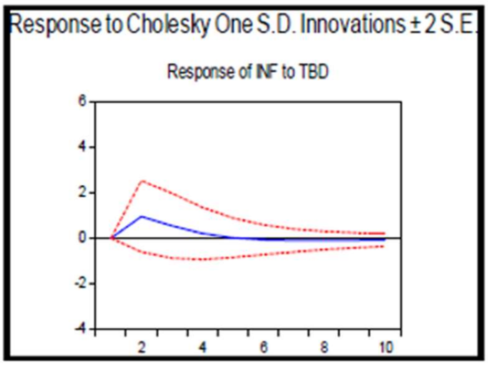

Fig. 6: Inflation Shock(Multivariate VAR)

Table 9: Comparison between the Impulse Shock Response of INF to TBD in the bivariate and multivariate VAR models

VAR Model / Comparison Aspect	Bivariate VAR Model (Fig. 5)	Multiple VAR Model (Fig. 6)
Initial Impact (Short-term Response)	Strong positive response, peaking around the second period. √	Positive response, peaking around the second period √ but less strong.
Statistical Significance (Short-term)	Statistically significant in the early periods. √	Statistically significant in the early periods √ but weaker.
Long-term Response	Gradually decreases and becomes insignificant after about 4 periods.	Gradually decreases and becomes insignificant after 4-5 periods.
Overall Trend	The effect fades over time and becomes insignificant, indicating the impact is temporary. √	Similar effect, √ but it lasts slightly longer before becoming insignificant.

Figure 7 shows the response of inflation (INF) to a one-standard-deviation shock in GDP and Figure 8 gives the response of inflation to a. one standard deviation in GEX.

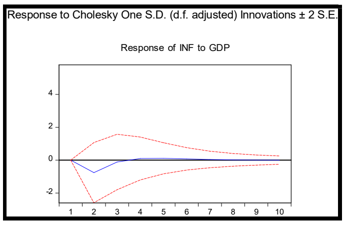

Fig. 7: INF to GDP

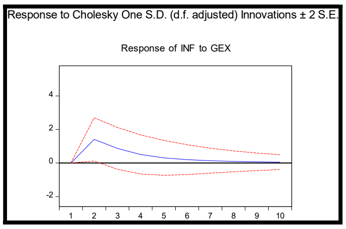

Fig. 8: INF to GEX

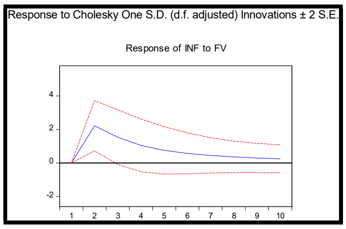

Fig. 9: INF to FV

In Figure 7. The short-term Response is slightly negative in the first period, suggesting that a shock to GDP reduces inflation in the short term. While the long-term Response is negative and gradually fades and becomes statistically insignificant after 3-4 periods, suggesting that the impact of GDP shocks on inflation disperses relatively quickly. In Figure 9, the short-term Response is positive and significant, with a sharp increase in the first period, indicating that financial shocks strongly increase inflation immediately. The Long-term Response fades after 3-4 periods, becoming insignificant, suggesting that the impact of foreign investments shocks on inflation is temporary and fades relatively quickly.

5. Conclusions and Recommendations

The current study aims to compare the bivariate VAR and the multivariate VAR with respect to model accuracy, model fit, variance decomposition and to conduct Impulse response analysis for some for some economic indicators of the Egyptian economy. EViews was used, variables used for the study are Inflation (INF), Trade Balance Deficit (TBD), Gross Domestic Product (GDP), Government Spending (GEX), and Foreign Investments (FV); data covered the period from 1990 to 2000. The following results were achieved:

1. The bivariate VAR model is preferred when analyzing the effect of a variable on itself, particularly for short-term shocks. For broader forecasting purposes, the multiple VAR model is more suitable.
2. Model Accuracy: The multiple VAR model significantly improves the explanatory power of inflation compared to the bivariate VAR model, with R^2 increasing from 29.19% (bivariate) to 50.39% (multiple).
3. Model Fit: The bivariate VAR model has lower AIC and SC values, indicating a simpler model, but the multiple VAR model provides better predictive performance despite the increased complexity.
4. Variance Contribution: In the multivariate VAR model, while inflation's self-contribution decreases, other variables like financial variables (FV) and government expenditure (GEX) take on more substantial explanatory roles.
5. Impulse Response Analysis: Both models show similar short-term responses to economic shocks, but the multivariate VAR model stabilizes inflation faster.
6. Long-term Stability: Stability tests confirm that the multivariate VAR model is more robust, indicating that including additional variables leads to better long-term forecasting of inflation.
7. Regular evaluation using statistical criteria like AIC and SC is crucial to balance model complexity and accuracy.

8. Detailed variance decomposition analysis when applying the multivariate VAR model captures the influence of multiple variables on the primary variable " inflation".
9. For forecast Accuracy, advanced statistical models, such as the multivariate VAR, should be employed to improve forecasting accuracy and enhance economic decision-making.

Recommendations.

1. The current study dealt with symmetric lag VAR models; it is recommended that a study on asymmetric VAR to be conducted.
2. Granger causality tests could be performed to find out if the relationship between "inflation and trade balance deficit", and "inflation and government spending" is bi-directional or uni-directional,
3. The impact of "Government spending" and " exchange rate: on "inflation" should be studied.
4. Variance decomposition and impulse shock response was studied only on "inflation"; it is recommended to extend the study to cover variance decomposition for each variable and to evaluate shock response for each variable on all others.

References

Agusti, B., & Costa, R. (2021). Economic shocks and variance decomposition in inflation forecasting: An application of multivariate VAR models. Journal of Applied Econometrics, 36(5), 978-995.

Aydin, A., & Cavdar, S. (2015). The impact of financial variables on inflation: Evidence from emerging markets. Journal of Economic Dynamics, 22(3), 345-361.

Bayraci, A., Demirhan, E., & Kilic, E. (2011). Forecasting inflation using dynamic factor models: Evidence from Turkey. Applied Economics, 43(6), 675-689.

Bose, E., Gupta, R., & Kabundi, A. (2017). Forecasting South African inflation using a dynamic factor model. South African Journal of Economics, 85(1), 49-67.

Braun, P. A., & Mittnik, S. (1993). Misspecifications in vector autoregressive models and their effects on impulse response and variance decomposition functions. Journal of Econometrics, 59(3), 319-341.

Davidson, R., & MacKinnon, J. G. (1993). Estimation and Inference in Econometrics. New York: Oxford University Press.

Feihan, Z., Wei, J., & Li, X. (2018). The impact of monetary policy shocks on inflation: Evidence from China. Economic Modelling, 69, 372-386.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37(3), 424-438.

Gujarati, D. N. (2009). Basic Econometrics (4th, 5th ed.). Military Academy, United States.

Hafer, R. W., & Sheehan, R. G. (1989). The sensitivity of VAR forecasts to alternative lag structures. International Journal of Forecasting, 5(3), 399-408.

Huang, Y., Liu, J., & Zhang, T. (2021). Inflation forecasting using machine learning: Evidence from G7 countries. Journal of Forecasting, 40(2), 253-270.

Kamal, A. (2023). A comparative study of inflation forecasting models: Application of machine learning techniques. Journal of Forecasting, 42(1), 45-62.

Li, Z., & Yuan, H. (2022). Inflation dynamics and monetary policy: New insights from VAR models. Economic Inquiry, 61(1), 34-58.

Longmore, P. (2020). Modeling inflation uncertainty using a multivariate GARCH model: Evidence from Caribbean economies. Journal of International Money and Finance, 102, 102129.

Lütkepohl, H. (1991). Introduction to Multiple Time Series Analysis. Berlin: Springer-Verlag.

Lütkepohl, H. (1999). Vector autoregressions. Unpublished manuscript, Institute for Statistics and Econometrics, Humboldt University at Berlin

Marcano, R. (2021). Monetary policy transmission mechanisms in Latin America: An empirical VAR analysis. Latin American Economic Review, 50(1), 87-104.

Musyoki, L., Mbutu, M., & Wanjiru, P. (2023). Predictive models for inflation: A VAR approach for African economies. African Journal of Economics, 10(1), 12-26.

Nalita, P., Siregar, R., & Yuniarti, D. (2021). Exchange rate pass-through and inflation: Evidence from Indonesia using VAR models. Indonesian Economic Review, 8(2), 67-84.

Nicholson, J., Glick, R., & Wang, P. (2022). How accurate are inflation forecasts? Evidence from advanced economies. International Economics, 170, 45-60.

Ozcicek, O., & McMillin, W. D. (1998). Lag length selection in vector autoregressive models: Symmetric versus asymmetric loss functions. Louisiana State University Working Paper.

Stock, J. H., & Watson, M. W. (2001). Vector autoregressions. Journal of Economic Perspectives, 15(4),

Waggoner, D. F., & Zha, T. (1998). Conditional forecasts in dynamic multivariate models. The Review of Economics and Statistics, 80(4), 639-651.

Watson, M. (1994). Vector autoregression and co-integration. In R.F. Engle & D.L. McFadden (Eds.), Handbook of Econometrics (Vol. 4, pp. 2843-2915). Elsevier.

Wei, S.-J. (2006). The effects of international trade on inflation: Evidence from OECD countries. Journal of International Economics, 67(1), 101-125.

Zivot, E., & Wang, J. (2006). Modeling Financial Time Series with S-PLUS. New York: Springer Science &Business Media.