Open Journal of Statistics, 2023, 13, 25-45
https://www.scirp.org/journal/ojs
ISSN Online: 2161-7198
ISSN Print: 2161-718X

Examine the Reliability of Econometrics Software: An Empirical Comparison of Time Series Modelling

Wickramasinghage M. A. Wickramasinghe[1*], Parana P. A. W. Athukorala[2], Siththara G. J. Senarathne[1], Yapa P. R. D. Yapa[1]

[1]Department of Statistics & Computer Science, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka
[2]Department of Economics & Statistics, Faculty of Arts, University of Peradeniya, Peradeniya, Sri Lanka
Email: *manojakalanka.maw@gmail.com, wathukorala@yahoo.com, jagath.senarathne@sci.pdn.ac.lk, roshany@sci.pdn.ac.lk

How to cite this paper: Wickramasinghe, W.M.A., Athukorala, P.P.A.W., Senarathne, S.G.J. and Yapa, Y.P.R.D. (2023) Examine the Reliability of Econometrics Software: An Empirical Comparison of Time Series Modelling. *Open Journal of Statistics*, **13**, 25-45.
https://doi.org/10.4236/ojs.2023.131003

Received: January 10, 2023
Accepted: February 19, 2023
Published: February 22, 2023

Abstract

Researchers must understand that naively relying on the reliability of statistical software packages may result in suboptimal, biased, or erroneous results, which affects applied economic theory and the conclusions and policy recommendations drawn from it. To create confidence in a result, several software packages should be applied to the same estimation problem. This study examines the results of three software packages (EViews, R, and Stata) in the analysis of time-series econometric data. The time-series data analysis which presents the determinants of macroeconomic growth of Sri Lanka from 1978 to 2020 has been used. The study focuses on testing for stationarity, cointegration, and significant relationships among the variables. The Augmented Dickey-Fuller and Phillips Perron tests were employed in this study to test for stationarity, while the Johansen cointegration test was utilized to test for cointegration. The study employs the vector error correction model to assess the short-run and long-term dynamics of the variables in an attempt to determine the relationship between them. Finally, the Granger Causality test is employed in order to examine the linear causation between the concerned variables. The study revealed that the results produced by three software packages for the same dataset and the same lag order vary significantly. This implies that time series econometrics results are sensitive to the software that is used by the researchers while providing different policy implications even for the same dataset. The present study highlights the necessity of further analysis to investigate the impact of software packages in time series analysis of economic scenarios.

Keywords

Econometrics, Macroeconomic Determinants, Software Packages, Time

Series Modelling

1. Introduction

Data analysis is a crucial part of research on which many researchers depend to make research results, conclusions, and recommendations. For a research project to be completed successfully, a rational data analysis procedure is vital. With the growing popularity of data analysis, a significant number of technological tools have developed in the academic context. Statistical software (SS) is one of these tools. The advent of SS has undoubtedly contributed significantly to the advancement of research studies in the 21st century.

Experts have shown that SS is a software application that simplifies the computation and presentation of data. If all data is entered properly, SS permits researchers to avoid frequent numerical errors and generate accurate results in their study. The development of SS enables scholarly researchers to quickly undertake more quantitative analyses. Many academics, experts, scientists, and business managers can also clearly convey accurate future predictions utilizing SS.

Modern statistical software has transformed statistical processes from time-consuming, arduous hand-calculation to working swiftly with large, complex datasets [1]. Incorporating statistical software into statistics education can assist students to enhance their statistical knowledge and enthusiasm toward statistics [2] [3] [4]. There are several proprietary and freeware SS programs available that are useful for various statistical analyses, depending on the demands of the user.

Researchers frequently attempt to uncover flaws in the statistical procedures, instead of treating the software as a possible source of error. Furthermore, many economists think that analyzing the reliability of software packages should be left to software developers, as the economics profession is less experienced in this area. According to McCullough & Vinod [5], this has made software flaws more frequent than they should be. Economists often place more importance on quickness and usability, but accuracy and reliability may suffer as a result. Moreover, they claim that as a result of these flaws, historical inferences are brought into question, and future work must document and archive statistical software in addition to statistical models.

Software testing research leads to better software. Software accuracy would be neglected if such research did not exist. The significant enhancements of statistical software show that developers are concerned about the reliability of their statistical software package. The fact that software developers, such as Stata and JMP, have posted their product's reliability on the web is a clear indication that vendors are paying attention. Stata Corporation rectified flaws reported in their benchmarking in response to an earlier conference-paper version of their study, according to Altman & McDonald [6]. While significant advances in computer technology have been achieved, it is still necessary to assess the reliability of statistical software packages on a regular basis to verify whether researchers are obtaining reliable and accurate estimates. According to studies, available software packages are not flawless, and may not be as efficient and accurate as some researchers believe [7] [8].

McCullough [9] evaluated versions 6.12, 7.5, and 4.0 of SAS, SPSS and S-Plus, particularly focusing on three areas: estimation (linear and nonlinear), random number generation, and statistical distribution. Although there were flaws in all three areas, the packages performed well when calculating univariate statistics and linear regression models. McCullough [10] examined four econometric statistical packages from a study of EViews, LIMDEP, SHAZAM and TSP and the author advised against relying on default nonlinear settings (e.g., algorithmic choice, tolerance level, starting point). Sawitzki [8] tested the numerical reliability of nine packages; BMDP, Data Desk, Excel, Interpretive Software Products (ISP), Generalized Linear Interactive Modelling (GLIM), SAS, Statistical Package for the Social Sciences (SPSS), S-Plus, and Statgraphics, and found that all of them failed basic computational tests like calculating sample variance estimators. McCullough & Heiser [11] investigated the accuracy of various statistical procedures in Excel 2007 and found that it failed accuracy tests for statistical distributions, random number generation, and estimation. L'ecuyer & Simard [12] created a software library for evaluating uniform random number generators and examined a series of frequently used random number generators. They discovered that the default generators in Excel, MATLAB, Mathematica, and R failed many of their tests. McCullough & Wilson [13] investigated the accuracy of statistical methods in Microsoft Excel 97. They found that Excel failed to perform well in linear and nonlinear estimates, random number generation, and statistical distributions (calculation of p-values). The authors advised against using Excel for statistical analyses. McCullough & Wilson [14] investigated the accuracy of statistical processes in Excel 2000 and Excel XP in a follow-up study. The authors noted that issues discovered by Sawitzki [8] in Excel 4.0, such as the instability of the sample variance estimator, were still present in Excel 2000 and Excel XP. McCullough & Wilson [7] indicated that, while there was some progress in some areas of Excel's computations, errors persisted in its updated versions. McCullough & Wilson [14] demonstrated in another study that issues highlighted in earlier versions of Excel were not resolved in Excel 2003. They noted that the software was still ineffective at handling nonlinear problems and warned users not to use it. McCullough [15] evaluated EViews 3.0, LIMDEP 7.0, RATS (Regression Analysis of Time Series) 4.3, SHAZAM 8.0, and TSP 4.4 for numerical accuracy. He found that, while the use of reliability testing led software developers to enhance these packages, numerous flaws still remained. He demanded that the updated versions of these packages be examined to see if the identified issues have been resolved. Keeling & Pavur [16] compared nine statistical software packages; SAS 9.1, SPSS 12.0, Excel 2003, Minitab 14.0, Stata 8.1, S-Plus 6.2, R 1.9.1, JMP 5.0, and StatCrunch 3.0 in the NIST benchmark tests with regard to univariate summary statistics, analysis of variance, linear regression, and nonlinear regression. They noted that when compared to earlier versions, newer versions of several applications like Excel 2003, SAS, SPSS, and S-Plus had made significant enhancements. They did, however, emphasize that errors remained in some of the packages.

According to the literature reviewed above, users should be aware of the limitations and shortcomings of different software packages. Furthermore, rather than depending on a single package, applying multiple packages and accepting a solution when the results agree across packages may be a more reliable approach. The useful contribution of this study is that it brings attention to the body of literature on variation in outputs by investigating software reliability across three statistical software packages (EViews, R and Stata) for econometric tests.

2. Methodology

Using annual time series data from 1978 to 2020, the study tries to identify the primary causes of growth for the Sri Lankan economy. The study's data came from the Special Statistical Appendix of the Central Bank of Sri Lanka's Annual Report 2020, which was accessed online. The real gross domestic product (*RGDP*), the exchange rate (*ER*), which signifies the exchange rates from USD to LKR, real public debt (*RPD*), real government expenditure (*RGE*), real imports (*RI*), real exports (*RE*), and broad money supply M2 (*BMS*) are among the variables considered in the study. The variables are all expressed in Sri Lankan rupees. Here, the variable *RGDP* is considered as a proxy to measure the economic growth of Sri Lanka.

$$RGDP = f\left(ER, RPD, RGE, RI, RE, BMS\right) \tag{1}$$

As such, the following Equation (2) is proposed to establish the relationship between economic growth and macroeconomic factors in Sri Lanka. According to the Equation (1), economic growth is affected by the exchange rate, real public debt, real government expenditure, real imports, real exports and broad money supply. Theory suggests that changes in the exchange rate, real public debt, real government expenditure, real imports, real exports and broad money supply will cause a change in real GDP. As a result of the aforementioned theoretical framework, the Sri Lankan economic growth function as it applies to this study may be stated in logarithmic linear form as follows:

$$\begin{aligned} \ln RGDP_t = \beta_0 + \beta_1 \ln ER_t + \beta_2 \ln RPD_t + \beta_3 \ln RGE_t + \beta_4 \ln RI_t \\ + \beta_5 \ln RE_t + \beta_6 \ln BMS_t + \varepsilon_t \end{aligned} \tag{2}$$

where $\ln RGDP_t$ represents the log of real GDP at time t;

$\ln ER_t$ represents the log of exchange rate at time t;

$\ln RPD_t$ represents the log of real public debt at time t;

$\ln RGE_t$ represents the log of real government expenditure at time t;

$\ln RI_t$ represents the log of real imports at time t;

$\ln RE_t$ represents the log of real exports at time t;

$\ln BMS_t$ represents the log of broad money supply at time t;

t = time

ε_t is the error term, which is assumed to be normally and independently distributed with a zero mean and constant variance which represents any other explanatory variables that impact economic development but are not represented by this model. In the model β_0 is a constant and β_1, β_2, β_3, β_4, β_5, β_6 denotes the sensitivities of each macroeconomic variable to economic growth. The study utilizes the VECM to evaluate the short-run and long-term dynamics of the variables in an attempt to identify the relationship between the variables. All preliminary tests for time series data were performed: test for stationarity using the ADF test and PP test, optimal lag length selection using the AIC, test for cointegration using the Johansen cointegration tests [17], later our model was fitted using the Johansen normalization technique. And finally, the Granger Causality test was carried out to determine the direction of causality between the variables. Analysis tests were performed using EViews 12.0, R 4.1.2, and Stata/SE 12.0 software.

2.1. Econometric Views (EViews)

EViews is a windows OS statistics tool that is mostly used for time-series related econometrics research. It was created by Quantitative Micro Software (QMS), which is now part of IHS. The most recent version of EViews is 12.0, which was published in November 2020. EViews can be used for general statistical and economic studies such as cross-section and panel data analysis, as well as time series estimation and forecasting. For data storage, EViews mainly depends on a proprietary and undocumented file format. However, it supports a wide range of input and output formats, including databank format, MS-Excel format, SPSS/PSPP, DAP/SAS, Stata, RATS, and TSP. Moreover, EViews can connect to Open Database Connectivity (ODBC) databases.

2.2. STATA

Stata is a robust statistical software that includes intelligent data-management features, a large range of up-to-date statistical methods, and an outstanding system for creating publication quality graphics. The most recent version Stata 17, released on April 20, 2021, is a quick and simple data management program. Stata is available on Windows, Unix, and Mac operating systems. Intercooled Stata (Stata/IC) is the standard version, and it can handle up to 2047 variables. Stata/SE is a special edition that can handle up to 32,766 variables (allows longer string variables and bigger matrices), while Stata/MP is a version for multicore/multiprocessor systems that has the same restrictions but is much faster. Stata conducts the majority of general statistical analyses including regression, logistic regression, survival analysis, analysis of variance, factor analysis, multivariate analysis and time series analysis.

2.3. R

R is a programming language for statistical computing developed by R development core team and the first official stable version was released in 2000. R is a free open-source software, which allows researchers to modify and develop packages as per their requirements. Another advantage of R is that R base environment can be run on multiple IDEs (RStudio, Jupyter notebook, etc.) based on the user preference. The comprehensive R archive network (CRAN) contains more than 10,000 different packages oriented at various different requirements. R also has special packages such as RShiny and ROrcale which enable users to build interactive webapps containing visualizations and interact with databases such as ROracle. The markdown package in R is also popular among users, as it helps to generate reports in the desired formats such as webpages, Word documents, PowerPoint slides, pdfs, etc. The latest version of R as of the date January 22, 2023 is the 4.2.2 version and at present it has grown to be one of the most popular statistical programming environments among researchers and analysts owing to the abovementioned distinguishing features.

3. Results and Discussion

Table 1 displays a descriptive statistic for each variable of interest. The average value of real GDP is Rs. 1,208,842 with the maximum and minimum values being Rs. 2,916,836 and Rs. 309,166 respectively during the study period. The average value of the real public debt is Rs. 1,028,986 with a maximum of Rs. 2,839,453 and a minimum of Rs. 221,713 and the dispersion from the mean value is Rs. 676,202. The average exchange rate (LKR to USD) is Rs. 78.63 while the maximum is Rs. 185.52 and the minimum is Rs. 15.57, with a wider dispersion from the mean of Rs. 49.07. Relative low value of the standard deviation of real export shows a less diversification of the export sector in the country. There is a significant growth of the import as well as money supply in the country during

Table 1. Summary statistics of variables used.

Variable	Mean	Maximum	Minimum	Std. Dev.
Real GDP (Rs.)	1208842.00	2916836.00	309166.70	840518.30
Exchange Rate (USD to LKR) (Rs.)	78.63	185.52	15.57	49.07
Real Public Debt (Rs.)	1028986.00	2839453.00	221713.00	676202.70
Real Expenditure (Rs.)	281863.50	648512.90	119950.00	141273.50
Real Import (Rs.)	5263.80	10852.91	3015.66	1521.34
Real Export (Rs.)	3451.58	6133.75	1887.19	978.26
Broad Money Supply (M2) (Rs.)	245847.60	1177150.00	5936.00	296472.60

this period. Having described the characteristics of the data and based on their background, we begin the econometric analysis by testing the order of integration using the ADF and PP unit root tests, and then proceed to test for cointegration using the Johansen Cointegration method, followed by testing for causality using the Granger Causality method [18].

3.1. Unit Root Tests

In this section the time-series properties of the data are examined during the period of 1978-2020, consequently ADF and PP tests are utilized. This testing is necessary to avoid spurious regression, which is a typical issue when estimating a regression line using data whose generated process follows a time trend. These unit-root tests were run in EViews, R and Stata on both levels and first differences of all variables. Furthermore, in order for the tests to be effective, the lag order must be chosen carefully so that the power of the test is not diminished [14]. Accordingly, the lag order was chosen using the minimum AIC values through EViews [19]. The R program takes the default value of $trunc\left\{\left(length\left(x\right)-1\right)\right\}^{\frac{1}{3}}$ as the lag order for ADF test, where x denotes the variable of interest. The Stata program allows the user to specify the number of lags to be used in the unit root tests. Both AIC and SBIC are often used to determine the optimal lag length. The aim is to choose the number of parameters that minimize the value of the information criterion. The SBIC has a propensity to underestimate the lag order, whereas adding more lags raises the penalty for the loss of degrees of freedom. Hence, AIC is used as the leading indicator to ensure that there is no residual autocorrelation in the VAR model. To study the discrepancies in the results, the three programs were run in the same lag order. **Table 2** shows the results of the ADF test performed by the three programs.

Table 2 indicates that for the three programs, all variables have a unit root at their level for the ADF test, and the p values for all series are not significant at the 1% and 5% significance levels. However, when the unit root test is run at the first difference, the ADF test results showed that all variables are stationary at the first difference, I (1), and the p-values are significant at 1% and 5% in all three programs. Based on these estimated results, we failed to reject the null hypothesis of unit roots at all level, but we reject the presence of unit root at all first differences of the variables. As a result, the variables are integrated of order 1, I (1). **Table 3** shows the results of the PP test performed by the three programs.

To test each series for unit roots, Phillips–Perron test utilized the Bartlett Kernel with automated Newey–West bandwidth selection. Similar to the ADF test results, the variables in the levels are nonstationary, but their first differences are stationary, I (1). Moreover, the test statistics for the ADF and PP tests in EViews and Stata are nearly identical, although the test statistics provided by the R application are distinct. The default lag value for the PP test in Stata is computed using $int\left(4\left(N/100\right)^{2/9}\right)$. To examine the variations in the outcomes, the same lag order is used to run the PP test in all three applications, as determined by the Bartlett Kernel with automatic Newey-West bandwidth selection through EViews. The results of the preceding unit-root tests reveal that all the seven variables are nonstationary in levels, but become stationary after taking their first difference, indicating that all the variables are integrated of order one, I (1) series. It is vital to emphasize that differencing a variable to attain stationarity might result in long run information loss. Thus, we perform a cointegration test to determine whether there is a long run relationship between the dependent and explanatory variables in various models.

Table 2. Augmented dickey-fuller unit root test for stationarity with intercept only.

Variables	EViews		R		Stata	
	Levels	First Difference	Levels	First Difference	Levels	First Difference
Log RGDP	−0.544 [0 lags] −0.872	−4.711 [0 lags] (0.000)***	−1.717 [0 lags] −0.686	−5.739 [0 lags] (0.000)***	−0.544 [0 lags] −0.883	−4.711 [0 lags] (0.000)***
Log Exchange Rate	−1.997 [0 lags] −0.287	−5.639 [0 lags] (0.000)***	−0.994 [0 lags] −0.928	−5.637 [0 lags] (0.000)***	−1.997 [0 lags] −0.288	−5.639 [0 lags] (0.000)***
Log Real Public Debt	−0.636 [0 lags] −0.852	−7.806 [0 lags] (0.000)***	−1.992 [0 lags] −0.577	−7.692 [0 lags] (0.000)***	−0.635 [0 lags] −0.863	−7.806 [0 lags] (0.000)***
Log Real Expenditure	1.929 [7 lags] −0.999	−5.55 [2 lags] (0.000)***	−1.405 [7 lags] −0.266	−3.483 [2 lags] (0.000)***	1.93 [7 lags] −0.999	−5.55 [2 lags] (0.000)***
Log Real Imports	−1.592 [0 lags] −0.478	−6.662 [1 lag] (0.000)***	−3.228 [0 lags] (0.096)*	−6.376 [1 lag] (0.000)***	−1.591 [0 lags] −0.488	−6.662 [1 lag] (0.000)***
Log Real Exports	−1.419 [0 lags] −0.564	−5.839 [1 lag] (0.000)***	−2.579 [0 lags] −0.345	−6.099 [1 lag] (0.000)***	−1.419 [0 lags] −0.573	−5.839 [1 lag] (0.000)***
Log Broad Money Supply (M2)	−2.335 [3 lags] −0.167	−4.347 [2 lags] (0.001)***	0.448 [3 lags] −0.99	−5.618 [2 lags] (0.000)***	−2.335 [3 lags] −0.161	−4.347 [2 lags] (0.000)***

a. MacKinnon (1996) one-sided p-values for rejection of hypothesis of unit root are denoted inside parentheses. *Denotes significance at the 10% level. **denotes significance at the 5% level. ***denotes significance at the 1% level.

Table 3. Phillips perron unit root test for stationarity with intercept only.

Variables	EViews		R		Stata	
	Levels	First Difference	Levels	First Difference	Levels	First Difference
Log RGDP	−0.515 [3] (−2.933)	−4.761 [2] (−2.935)***	0.129 [3] (−2.932)	−5.832 [2] (−2.934)***	−0.514 [3] (−2.952)***	−4.761 [2] (−2.955)***
Log Exchange Rate	−2.033 [2] (−2.933)	−5.719 [4] (−2.935)***	1.619 [2] (−2.932)	−5.421 [4] (−2.934)***	−2.033 [2] (−2.952)***	−5.719 [4] (−2.955)***

Continued

Log Real Public Debt	−0.651 [3] (−2.933)	−7.806 [0] (−2.935)***	0.156 [3] (−2.932)	−7.679 [0] (−2.934)***	−0.651 [3] (−2.952)***	−7.806 [0] (−2.955)***
Log Real Expenditure	1.102 [41] (−2.933)	−18.034 [40] (−2.935)***	−1.660 [41] (−2.932)	−14.479 [40] (−2.934)***	1.102 [41] (−2.952)***	−18.034 [40] (−2.955)***
Log Real Imports	−1.215 [4] (−2.933)	−7.088 [5] (−2.935)***	−2.208 [4] (−2.932)	−7.345 [5] (−2.934)***	−1.215 [4] (−2.952)***	−7.088 [5] (−2.955)***
Log Real Exports	−1.419 [0] (−2.933)	−6.538 [0] (−2.935)***	−1.076 [0] (−2.932)	−7.055 [0] (−2.934)***	−1.419 [0] (−2.952)***	−6.538 [0] (−2.955)***
Log Broad Money Supply (M2)	−3.069 [24] (−2.933)**	−5.238 [40] (−2.935)***	2.957 [24] (−2.932)	−5.56 [40] (−2.934) ***	−3.069 [24] (−2.952)**	−5.238 [40] (−2.955)***

a. MacKinnon (1996) critical values for rejection of hypothesis of unit root are denoted in parentheses. *Denotes significance at the 10% level. **denotes significance at the 5% level. ***denotes significance at the 1% level. The spectral estimation is done using the Bartlett kernel with Newey–West automatic bandwidth. Bandwidth is reported inside square brackets.

3.2. Cointegration Tests

The Johansen Maximum cointegration test [17] is applied based on the above unit-root tests. Before conducting cointegration tests, the relevant order of lags (p) of the VAR model must be specified. Based on the minimum AIC values, EViews suggests lag order 2 for the VAR model. The same lag order is utilized in R and Stata software to conduct cointegration tests. However, in the R application, the information criteria specified lag order 5 under the VAR select function. EViews and Stata software failed to produce acceptable results for the Johansen cointegration test at lag 5. Furthermore, despite lag selection using the R application, the first two test statistics of $r = 0$ and $r \leq 1$ were generated as NaN. Consequently, lag 2 was used in all three applications for the test.

The results of the maximum eigenvalue and trace tests are shown in **Table 4**, starting with the null hypothesis that there is no cointegration ($r = 0$) among the seven variables. The maximum eigenvalue and the trace statistic in both EViews and R suggest $r = 2$. As a result, it is inferred that the variables have two cointegrating relationships. In Stata, the maximum eigenvalue and trace statistic indicate $r = 1$, implying that there is only one cointegrating relationship among the variables. This result clearly shows that different statistical programs generate the different results for econometric modelling of macroeconomic variables. **Table 5** shows estimates of long-run cointegrating vectors.

Empirical results of the long run relationship taken from EViews, R and Stata with their significance level are summarized in **Table 5**. Interestingly, R application does not provide the information that are needed to decide significance of the variables. According to Engle & Granger [20], Marquez [21] testing the level of significance of the standardized cointegration vector is not required.

Table 4. Johansen cointegration tests.

Software	No. of CE (s)	Trace Statistic	5% Critical Value	Maximum	5% Critical Value
EViews	None*	170.533	125.615	56.387	46.231
	At most 1*	114.146	95.754	54.127	40.078
	At most 2	60.019	69.819	24.679	33.877
	At most 3	35.34	47.856	14.242	27.584
	At most 4	21.098	29.797	10.695	21.132
	At most 5	10.404	15.495	8.838	14.265
	At most 6	1.566	3.841	1.566	3.841
R	None*	164.87	131.7	48.81	46.45
	At most 1*	116.06	102.14	45.08	40.3
	At most 2	70.98	76.07	26.14	34.4
	At most 3	44.84	53.12	17.62	28.14
	At most 4	27.23	34.91	12.52	22
	At most 5	14.7	19.96	9.9	15.67
	At most 6	4.81	9.24	4.81	9.24
Stata	None*	138.437	124.24	46.619	45.28
	At most 1	91.818	94.15	34.725	39.37
	At most 2	57.092	68.52	26.004	33.46
	At most 3	31.088	47.21	15.181	27.07
	At most 4	15.908	29.68	9.472	20.97
	At most 5	6.436	15.41	4.784	14.07
	At most 6	1.652	3.76	1.652	3.76

a. *denotes rejection of hypothesis at the 0.05 level.

Table 5. Normalized cointegration vectors.

Variables	EViews		R	Stata
Log RGDP	−0.072	0	−0.999	−0.891***
Log Exchange Rate	3.120***	2.894***	2.257	1.892***
Log Real Public Debt	−3.853***	−3.696***	−0.931	−0.599***
Log Real Expenditure	1.871***	1.827***	−0.629	−0.745***
Log Real Imports	−1.462***	−1.447***	0.842	0.928***
Log Real Exports	−0.687***	−0.669**	−0.708	−0.694***
Log Broad Money Supply (M2)	−4.285***	-	−22.062	−21.242***

a. *Denotes significance at the 10% level. **denotes significance at the 5% level; ***denotes significance at the 1% level.

The EViews and R software generated two normalized cointegration vectors, whereas Stata generated one normalized cointegration vector. In each of the three applications, there is a distinct variation in the coefficients of the long run relationship. The long run models can be re-parameterized as follows.

Table 6 demonstrates a significant difference in the long run models generated by the three programs. Despite the fact that the Johansen cointegration results of R program revealed the existence of two cointegration equations, the program produced only one normalized cointegration vector for the variable *RGDP*. The signs of the coefficients of the variables in the long run models created by R and Stata are identical, which is a distinguishing feature of the long run models developed by normalizing for real GDP variable; however, the signs of the variables *lnRI* and *lnRE* are opposite in EViews when compared with the other two programs. Moreover, coefficient of the *lnRE* variable produced by EViews is 1.462 whereas the coefficient produced by R and Stata are −0.842 and −0.928 respectively. The coefficients produced for *lnRPD* are −3.12, −2.257 and −1.892 by EViews, R and Stata respectively, suggesting a negative impact from real public debt towards real GDP in the long run. The coefficients produced by EViews, R, and Stata for *lnER* are 0.072, 0.999, and 0.891 respectively, indicating that the exchange rate has a positive impact on real GDP in the long run. The existence of cointegration necessitates the estimation of an ECM, which represents the short-run dynamics. Therefore, in order to fully investigate the impact of macroeconomic variables on economic growth, short-run estimations were performed in this paper that is discussed in the next section.

Table 6. Long run models of the three software.

EViews	lnRGDP = 4.285 + 0.072lnER − 3.12lnRPD + 3.853lnRGE − 1.871lnRI + 1.462lnRE + 0.687lnBMS
R	lnRGDP = 22.062 + 0.999lnER − 2.257lnRPD + 0.931lnRGE + 0.629lnRI − 0.842lnRE + 0.708lnBMS
Stata	lnRGDP = 21.242 + 0.891lnER − 1.892lnRPD + 0.599lnRGE + 0.745lnRI − 0.928lnRE + 0.694lnBMS

3.3. Vector Error Correction Model (VECM)

The VECM is used to investigate the short run dynamics of the variables. The error correction model allows previous disequilibrium to be introduced as independent variables in the dynamic behavior of existing variables. VECM relates changes in RGDP growth to changes in the other lagged variables and the disturbance term of lagged periods. Table 7 shows the results of vector error correction model's short run dynamic relationship and set of short run coefficients for the three software.

The EViews results show that the coefficient of the ECT is negative and is not significant at 5 percent. Additionally, the immediate impact of the explanatory variables demonstrates that the last two years of real public debt and real exports had negative impacts on the growth of real GDP and the last year of real government expenditure and broad money supply M2 had negative and positive impacts on real GDP respectively. These impacts were statistically significant. Therefore, an increase in the past two years of real public debt will cause growth in real GDP to decline by 0.24 percent while an increase in the past year of broad money supply will cause growth in real GDP to increase by 0.35 percent. Moreover, an increase in the past year of real government expenditure will cause the growth of real GDP to fall by 0.27 percent. Alternatively, the past two-year records of growth in real GDP had a significant positive impact on current growth in real GDP. Therefore, in the short run, real public debt, real government expenditure, real exports and broad money supply are determinants of growth of real GDP.

In contrast, the R results show that the coefficient of the ECT is negative and significant at 1 percent. This implies that each year, a 15.8 percentage point modification is made to the long run periods. Table 7 also indicates that the past year of broad money supply had a negative significant impact on growth of real GDP. Therefore, an increase in the broad money supply in the past year will cause the growth of real GDP to decrease by 0.3 percent. The past records of exchange rate, real government expenditure and real exports had a negative impact on growth in real GDP. Moreover, the past records of real imports had a positive impact on the growth of real GDP. However, the impacts of these variables are not statistically significant.

The coefficient of the EC term is negative and significant at 1 percent, according to the Stata results. This means that the long run periods are modified by 24.8 percentage points per year. The past year of real public debt, real exports and broad money supply had a positive significant impact on the growth of real GDP. Furthermore, records of real government expenditure and real imports in the past year had a significant negative impact on the growth of real GDP. Therefore, an increase in the real public debt in the past year will cause the growth in real GDP to increase by 0.34 percent. An increase in the past year of real government expenditure will cause the growth in real GDP to decline by 0.25 percent. Whereas a rise in real imports in the previous year caused a 0.11 percent drop in real GDP growth. Moreover, an increase in real exports over the last year will result in a 0.18 growth in real GDP. Hence, in the short run model produced by Stata, real public debt, real government expenditure, real imports, real exports and broad money supply are determinants of growth of real GDP.

Although the co-integration test helps detect the long-term equilibrium relationship between variables, it cannot identify the causal relationship between variables. Therefore, Granger causality test is employed to determine the causal relationship between the variables.

Table 7. VECM Estimation for *lnRGDP*.

Dependent variable: Δ(Log RGDP)			
41 observations used for estimation from 1980 to 2020			
Regressors	EViews	R	Stata
	Parameter estimates (Std. Error)	Parameter estimates (Std. Error)	Parameter estimates (Std. Error)
Constant	0.019 (0.024)	−0.814 (0.247)***	0.101 (0.019)***
Δ(Log RGDP (−1))	−0.039 (0.175)	−0.026 (0.192)	−0.258 (0.167)
Δ(Log RGDP (−2))	0.461 (0.172)**	0.355 (0.182)	-
Δ(Log Exchange Rate (−1))	0.049 (0.176)	−0.136 (0.169)	0.055 (0.132)
Δ(Log Exchange Rate (−2))	−0.197 (0.168)	−0.021 (0.194)	-
Δ(Log Real Public Debt (−1))	−0.055 (0.118)	0.013 (0.111)	0.342 (0.106)***
Δ(Log Real Public Debt (−2))	−0.244 (0.106)**	−0.046 (0.112)	-
Δ(Log Real Expenditure (−1))	−0.276 (0.121)**	−0.078 (0.101)	−0.249 (0.065)***
Δ(Log Real Expenditure (−2))	−0.089 (0.107)	−0.086 (0.084)	-
Δ(Log Real Imports (−1))	−0.012 (0.069)	0.122 (0.080)	−0.114 (0.047)***
Δ(Log Real Imports (−2))	0.029 (0.067)	0.060 (0.082)	-
Δ(Log Real Exports (−1))	−0.125 (0.076)	−0.058 (0.097)	0.180 (0.077)***
Δ(Log Real Exports (−2))	−0.181 (0.079)**	−0.127 (0.085)	-
Δ(Log Broad Money Supply (M2) (−1))	0.351 (0.103)***	−0.301 (0.118)**	0.322 (0.067)***
Δ(Log Broad Money Supply (M2) (−2))	0.011 (0.101)	−0.110 (0.095)	-
ECT	−0.049 (0.027)*	−0.158 (0.052)***	−0.248 (0.047)***

a. *Denotes significance at the 10% level. **denotes significance at the 5% level; ***denotes significance at the 1% level.

3.4. Granger Causality Test

Granger [22] causality test has been performed in order to examine the linear causation between the concerned variables. Granger causality is useful in determining the direction of the relationships. The presence of a cointegration vector, according to Granger, indicates that Granger causality must occur in at least one direction. Tables 8-10 exhibit the results of the Granger causality test conducted by EViews, Stata, and R, respectively, using the optimal lag length of 1 based on AIC.

From Table 8, the result shows that there is unilateral directional causality between real GDP and real imports, exchange rate and real government expenditure, real imports and exchange rate, real government expenditure and real imports, real exports and real government expenditure. Also, there is bidirectional causality between broad money supply and real imports.

Table 9 shows that there is unilateral directional causality between real public debt and real GDP, real government expenditure and real GDP, real imports and real GDP, real exports and real GDP. Table 10 shows unilateral directional causality for real public debt and real GDP, real GDP and real imports, real GDP and real exports, exchange rate and real public debt, exchange rate and real imports, exchange rate and real exports, and bidirectional causality between real GDP and real government expenditure. Accordingly, it can be concluded that there is a significant variation of the results produced by different statistical packages in time series modelling in the field of macroeconomics. The results generated by econometric analyses should be interpreted and explained in the context of economic theory and principles [4]. This will lead bias results and any policies drawn from the results taken from using different computer packages may mislead the true behavior of the macroeconomic variables.

Table 8. Granger causality test results EViews.

Null hypothesis	F-Statistic	p-value	Decision
D (Log Exchange Rate) does not Granger Cause D (Log RGDP)	0.297	0.589	Not Reject
D (Log RGDP) does not Granger Cause D (Log Exchange Rate)	1.407	0.243	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log RGDP)	0.037	0.848	Not Reject
D (Log RGDP) does not Granger Cause D (Log Real Public Debt)	0.002	0.965	Not Reject
D (Log Real Expenditure) does not Granger Cause D (Log RGDP)	0.209	0.65	Not Reject
D (Log RGDP) does not Granger Cause D (Log Real Expenditure)	0.779	0.383	Not Reject
D (Log Real Imports) does not Granger Cause D (Log RGDP)	2.585	0.116	Not Reject
D (Log RGDP) does not Granger Cause D (Log Real Imports)	6.444	0.015	Reject
D (Log Real Export) does not Granger Cause D (Log RGDP)	3.121	0.085	Not Reject
D (Log RGDP) does not Granger Cause D (Log Real Export)	1.534	0.223	Not Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log RGDP)	4.684	0.037	Reject
D (Log RGDP) does not Granger Cause D (Log Broad Money Supply M2)	3.539	0.068	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log Exchange Rate)	2.187	0.148	Not Reject
D (Log Exchange Rate) does not Granger Cause D (Log Real Public Debt)	0.842	0.365	Not Reject
D (Log Real Expenditure) does not Granger Cause D (Log Exchange Rate)	0.801	0.376	Not Reject
D (Log Exchange Rate) does not Granger Cause D (Log Real Expenditure)	6.328	0.016	Reject
D (Log Real Imports) does not Granger Cause D (Log Exchange Rate)	10.103	0.003	Reject
D (Log Exchange Rate) does not Granger Cause D (Log Real Imports)	3.717	0.061	Not Reject
D (Log Real Export) does not Granger Cause D (Log Exchange Rate)	1.731	0.196	Not Reject
D (Log Exchange Rate) does not Granger Cause D (Log Real Export)	0.36	0.552	Not Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log Exchange Rate)	0.057	0.812	Not Reject
D (Log Exchange Rate) does not Granger Cause D (Log Broad Money Supply M2)	0.386	0.538	Not Reject
D (Log Real Expenditure) does not Granger Cause D (Log Real Public Debt)	0.114	0.738	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log Real Expenditure)	2.978	0.093	Not Reject
D (Log Real Imports) does not Granger Cause D (Log Real Public Debt)	0.335	0.566	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log Real Imports)	3.298	0.077	Not Reject

Continued

D (Log Real Export) does not Granger Cause D (Log Real Public Debt)	0.027	0.869	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log Real Export)	0.457	0.503	Not Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log Real Public Debt)	0.182	0.672	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log Broad Money Supply M2)	0.02	0.888	Not Reject
D (Log Real Imports) does not Granger Cause D (Log Real Expenditure)	2.808	0.102	Not Reject
D (Log Real Expenditure) does not Granger Cause D (Log Real Imports)	4.389	0.043	Reject
D (Log Real Export) does not Granger Cause D (Log Real Expenditure)	5.669	0.022	Reject
D (Log Real Expenditure) does not Granger Cause D (Log Real Export)	2.966	0.093	Not Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log Real Expenditure)	5.485	0.025	Reject
D (Log Real Expenditure) does not Granger Cause D (Log Broad Money Supply M2)	0.976	0.329	Not Reject
D (Log Real Export) does not Granger Cause D (Log Real Imports)	2.226	0.144	Not Reject
D (Log Real Imports) does not Granger Cause D (Log Real Export)	1.559	0.219	Not Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log Real Imports)	4.118	0.049	Reject
D (Log Real Imports) does not Granger Cause D (Log Broad Money Supply M2)	5.127	0.029	Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log Real Export)	2.767	0.104	Not Reject
D (Log Real Export) does not Granger Cause D (Log Broad Money Supply M2)	1.085	0.304	Not Reject

Table 9. Granger causality test results Stata.

Null hypothesis	Chi-Statistic	p-value	Decision
D (Log Exchange Rate) does not Granger Cause D (Log RGDP)	0.18	0.675	Not Reject
D (Log RGDP) does not Granger Cause D (Log Exchange Rate)	2.58	0.1082	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log RGDP)	10.34	0.0013	Reject
D (Log RGDP) does not Granger Cause D (Log Real Public Debt)	0.31	0.5802	Not Reject
D (Log Real Expenditure) does not Granger Cause D (Log RGDP)	14.85	0.0001	Reject
D (Log RGDP) does not Granger Cause D (Log Real Expenditure)	0.75	0.3866	Not Reject
D (Log Real Imports) does not Granger Cause D (Log RGDP)	5.93	0.0149	Reject
D (Log RGDP) does not Granger Cause D (Log Real Imports)	1.9	0.1686	Not Reject
D (Log Real Export) does not Granger Cause D (Log RGDP)	5.45	0.0196	Reject
D (Log RGDP) does not Granger Cause D (Log Real Export)	0.54	0.4641	Not Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log RGDP)	23.19	0	Reject
D (Log RGDP) does not Granger Cause D (Log Broad Money Supply M2)	0	0.9608	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log Exchange Rate)	0.12	0.7258	Not Reject
D (Log Exchange Rate) does not Granger Cause D (Log Real Public Debt)	1.82	0.177	Not Reject
D (Log Real Expenditure) does not Granger Cause D (Log Exchange Rate)	0.07	0.7878	Not Reject
D (Log Exchange Rate) does not Granger Cause D (Log Real Expenditure)	0.4	0.5285	Not Reject

Continued

D (Log Real Imports) does not Granger Cause D (Log Exchange Rate)	6.89	0.0087	Reject
D (Log Exchange Rate) does not Granger Cause D (Log Real Imports)	1.98	0.1599	Not Reject
D (Log Real Export) does not Granger Cause D (Log Exchange Rate)	0.05	0.8292	Not Reject
D (Log Exchange Rate) does not Granger Cause D (Log Real Export)	1.21	0.2705	Not Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log Exchange Rate)	0.03	0.8652	Not Reject
D (Log Exchange Rate) does not Granger Cause D (Log Broad Money Supply M2)	3.89	0.0487	Reject
D (Log Real Expenditure) does not Granger Cause D (Log Real Public Debt)	0.11	0.736	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log Real Expenditure)	0.01	0.938	Not Reject
D (Log Real Imports) does not Granger Cause D (Log Real Public Debt)	0.1	0.7478	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log Real Imports)	0.05	0.8303	Not Reject
D (Log Real Export) does not Granger Cause D (Log Real Public Debt)	1.3	0.2544	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log Real Export)	1.1	0.2943	Not Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log Real Public Debt)	0.11	0.7406	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log Broad Money Supply M2)	0.46	0.4953	Not Reject
D (Log Real Imports) does not Granger Cause D (Log Real Expenditure)	0.08	0.7806	Not Reject
D (Log Real Expenditure) does not Granger Cause D (Log Real Imports)	4.53	0.0333	Reject
D (Log Real Export) does not Granger Cause D (Log Real Expenditure)	1.73	0.1884	Not Reject
D (Log Real Expenditure) does not Granger Cause D (Log Real Export)	4.39	0.0361	Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log Real Expenditure)	4.38	0.0363	Reject
D (Log Real Expenditure) does not Granger Cause D (Log Broad Money Supply M2)	0.01	0.934	Not Reject
D (Log Real Export) does not Granger Cause D (Log Real Imports)	0.82	0.3661	Not Reject
D (Log Real Imports) does not Granger Cause D (Log Real Export)	5.34	0.0209	Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log Real Imports)	4.63	0.0314	Reject
D (Log Real Imports) does not Granger Cause D (Log Broad Money Supply M2)	3.29	0.0698	Not Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log Real Export)	3.95	0.0469	Reject
D (Log Real Export) does not Granger Cause D (Log Broad Money Supply M2)	0.46	0.4989	Not Reject

Table 10. Granger causality test results R.

Null hypothesis	F-Statistic	p-value	Decision
D (Log Exchange Rate) does not Granger Cause D (Log RGDP)	0.4277	0.515	Not Reject
D (Log RGDP) does not Granger Cause D (Log Exchange Rate)	1.0543	0.3077	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log RGDP)	9.1643	0.0033	Reject
D (Log RGDP) does not Granger Cause D (Log Real Public Debt)	0.3809	0.5389	Not Reject
D (Log Real Expenditure) does not Granger Cause D (Log RGDP)	6.0761	0.0159	Reject
D (Log RGDP) does not Granger Cause D (Log Real Expenditure)	13.263	0.0005	Reject

Continued

D (Log Real Imports) does not Granger Cause D (Log RGDP)	2.4624	0.1206	Not Reject
D (Log RGDP) does not Granger Cause D (Log Real Imports)	4.5061	0.0369	Reject
D (Log Real Export) does not Granger Cause D (Log RGDP)	2.492	0.1185	Not Reject
D (Log RGDP) does not Granger Cause D (Log Real Export)	6.4628	0.013	Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log RGDP)	0.6526	0.4217	Not Reject
D (Log RGDP) does not Granger Cause D (Log Broad Money Supply M2)	0.3686	0.5455	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log Exchange Rate)	0.0313	0.86	Not Reject
D (Log Exchange Rate) does not Granger Cause D (Log Real Public Debt)	11.272	0.0012	Reject
D (Log Real Expenditure) does not Granger Cause D (Log Exchange Rate)	0.4029	0.5274	Not Reject
D (Log Exchange Rate) does not Granger Cause D (Log Real Expenditure)	3.2622	0.0747	Not Reject
D (Log Real Imports) does not Granger Cause D (Log Exchange Rate)	0.9508	0.3325	Not Reject
D (Log Exchange Rate) does not Granger Cause D (Log Real Imports)	6.8212	0.0108	Reject
D (Log Real Export) does not Granger Cause D (Log Exchange Rate)	0.1716	0.6798	Not Reject
D (Log Exchange Rate) does not Granger Cause D (Log Real Export)	4.6468	0.0342	Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log Exchange Rate)	3.6331	0.0603	Not Reject
D (Log Exchange Rate) does not Granger Cause D (Log Broad Money Supply M2)	3.442	0.0673	Not Reject
D (Log Real Expenditure) does not Granger Cause D (Log Real Public Debt)	0.9777	0.3258	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log Real Expenditure)	6.2934	0.0142	Reject
D (Log Real Imports) does not Granger Cause D (Log Real Public Debt)	2.0745	0.1538	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log Real Imports)	5.7311	0.0191	Reject
D (Log Real Export) does not Granger Cause D (Log Real Public Debt)	0.5726	0.4515	Not Reject
D (Log Real Public Debt) does not Granger Cause D (Log Real Export)	5.7422	0.0189	Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log Real Public Debt)	10.265	0.0019	Reject
D (Log Real Public Debt) does not Granger Cause D (Log Broad Money Supply M2)	0.1599	0.6904	Not Reject
D (Log Real Imports) does not Granger Cause D (Log Real Expenditure)	2.8436	0.0957	Not Reject
D (Log Real Expenditure) does not Granger Cause D (Log Real Imports)	3.6664	0.0592	Not Reject
D (Log Real Export) does not Granger Cause D (Log Real Expenditure)	3.3637	0.0705	Not Reject
D (Log Real Expenditure) does not Granger Cause D (Log Real Export)	3.7345	0.0569	Not Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log Real Expenditure)	5.8222	0.0182	Reject
D (Log Real Expenditure) does not Granger Cause D (Log Broad Money Supply M2)	0.0007	0.9783	Not Reject
D (Log Real Export) does not Granger Cause D (Log Real Imports)	0.0664	0.7974	Not Reject
D (Log Real Imports) does not Granger Cause D (Log Real Export)	0.8275	0.3658	Not Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log Real Imports)	6.9644	0.01	Reject
D (Log Real Imports) does not Granger Cause D (Log Broad Money Supply M2)	0.8431	0.3613	Not Reject
D (Log Broad Money Supply M2) does not Granger Cause D (Log Real Export)	7.3229	0.0084	Reject
D (Log Real Export) does not Granger Cause D (Log Broad Money Supply M2)	1.5458	0.2175	Not Reject

4. Conclusions and Future Work

The primary goal of this study is to examine the discrepancies in the results of econometric analysis tests performed by EViews, R and Stata on the same data-set. The procedures used in all three programs are similar and the same data set is used. To test the stationarity and the order of integration of all series, the ADF test and the PP test were employed and Akaike Information Criterion was used to find the optimal lag structure for each variable. The results of these tests indicated that all the variables used in the study are integrated into order 1. The test statistics of the ADF test and PP test produced by EViews and Stata programs are nearly identical for the same lag order and the test statistic of the tests produced by R are distinct when compared with EViews and Stata. Later an econometric analysis was conducted based on the Johansen cointegration test and a vector error correction model to analyze the interrelationships between macroeconomic variables and economic growth. The Johansen cointegration tests carried out in EViews and R programs inferred that there exists a long run equilibrium relationship between the variables and the presence of two cointegrating equations at a 0.05 significance level. The Johansen test procedure run in the Stata program confirmed that there is at least one cointegrating equation at 0.05 significance for the economic growth and macroeconomic variables. The coefficients of the long run model produced by the Johansen test in the three programs show a significant variation when compared with each other. The signs of the variables in the long-run model produced by R and Stata are identical when compared with the EViews program. In the long run, economic growth is positively impacted by exchange rate, real government expenditure and broad money supply, while real public debt negatively impacts the economic growth. In the R and Stata programs, real imports and real exports had a positive and negative impact on economic growth, respectively. Whereas in EViews, the impact of real imports and real exports were negative and positive on economic growth, respectively.

To assess the short run dynamics of the variables vector error correction mechanism was carried out and the results of the VEC in EViews show that short run relationships exist between real GDP, real public debt, real government expenditure, real exports and broad money supply. The R results of VEC show that there exists a short-run relationship between real GDP and broad money supply. The VEC run in the Stata program indicated that there exist significant short run relationships between real GDP, real public debt, real government expenditure, real imports, real exports and broad money supply. The Granger causality test findings also revealed differences in the three software. EViews and R showed that there is a unilateral directional causality between real GDP and real imports. The results of EViews and Stata yielded a similar conclusion for unilateral directional causation between broad money supply and real GDP. For unilateral directional causations between broad money supply with real imports, real exports and real government expenditure, the Stata and R findings yielded the same conclusion.

Taken altogether, these findings suggest that the statistical econometric analyses performed by different applications yield diverse results. This paper provides new insights related to the results of tests performed in different applications. The main explanation for the disparities in results might be changes in the theoretical formulations of the tests listed. More investigation should be conducted to figure out the actual cause of these disparities in results across different applications.

Acknowledgements

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Chen, E.E. and Wojcik, S.P. (2016) A Practical Guide to Big Data Research in Psychology. *Psychological Methods*, **21**, 458-474. https://doi.org/10.1037/met0000111

[2] Ciftci, S.K., Karadag, E. and Akdal, P. (2014) Instruction of Statistics via Computer-Based Tools: Effects on Statistics' Anxiety, Attitude, and Achievement. *Journal of Educational Computing Research*, **50**, 119-133. https://doi.org/10.2190/EC.50.1.f

[3] McCulloch, R S. (2017) Learning Outcomes in a Laboratory Environment vs. Classroom for Statistics Instruction: An Alternative Approach Using Statistical Software. *International Journal of Higher Education*, **6**, 131-142. https://doi.org/10.5430/ijhe.v6n5p131

[4] Sosa, G.W., Berger, D.E., Saw, A.T. and Mary, J.C. (2011) Effectiveness of Computer-Assisted Instruction in Statistics: A Meta-Analysis. *Review of Educational Research*, **81**, 97-128. https://doi.org/10.3102/0034654310378174

[5] McCullough, B.D. and Vinod, H.D. (1999) The Numerical Reliability of Econometric Software. *Journal of Economic Literature*, **37**, 633-665. https://doi.org/10.1257/jel.37.2.633

[6] Altman, M. and McDonald, M.P. (2001) Choosing Reliable Statistical Software. *PS: Political Science & Politics*, **34**, 681-687. https://doi.org/10.1017/S1049096501001093

[7] McCullough, B.D. and Wilson, B. (2005) On the Accuracy of Statistical Procedures in Microsoft Excel 2003. *Computational Statistics & Data Analysis*, **49**, 1244-1252. https://doi.org/10.1016/j.csda.2004.06.016

[8] Sawitzki, G. (1994) Report on the Numerical Reliability of Data Analysis Systems. *Computational Statistics & Data Analysis*, **18**, 289-301. https://doi.org/10.1016/0167-9473(94)90177-5

[9] McCullough, B.D. (1999) Assessing the Reliability of Statistical Software: Part II. *The American Statistician*, **53**, 149-159. https://doi.org/10.1080/00031305.1999.10474450

[10] McCullough, B.D. (1999) Econometric Software Reliability: EViews, LIMDEP,

SHAZAM and TSP. *Journal of Applied Econometrics*, **14**, 191-202. https://doi.org/10.1002/(SICI)1099-1255(199903/04)14:2<191::AID-JAE524>3.0.CO;2-K

[11] McCullough, B.D. and Heiser, D.A. (2008) On the Accuracy of Statistical Procedures in Microsoft Excel 2007. *Computational Statistics & Data Analysis*, **52**, 4570-4578. https://doi.org/10.1016/j.csda.2008.03.004

[12] L'Ecuyer, P. and Simard, R. (2007) TestU01: AC Library for Empirical Testing of Random Number Generators. *ACM Transactions on Mathematical Software* (*TOMS*), **33**, Article No. 22. https://doi.org/10.1145/1268776.1268777

[13] McCullough, B.D. and Wilson, B. (2002) On the accuracy of statistical procedures in Microsoft Excel 2000 and Excel XP. *Computational Statistics & Data Analysis*, **40**, 713-721. https://doi.org/10.1016/S0167-9473(02)00095-6

[14] Ng, S. and Perron, P. (2001) LAG Length Selection and the Construction of Unit Root Tests with Good Size and Power. *Econometrica*, **69**, 1519-1554. https://doi.org/10.1111/1468-0262.00256

[15] McCullough, B.D. (2004) Wilkinson's Tests and Econometric Software. *Journal of Economic and Social Measurement*, **29**, 261-270. https://doi.org/10.3233/JEM-2004-0199

[16] Keeling, K.B. and Pavur, R.J. (2007) A Comparative Study of the Reliability of Nine Statistical Software Packages. *Computational Statistics & Data Analysis*, **51**, 3811-3831. https://doi.org/10.1016/j.csda.2006.02.013

[17] Johansen, S. and Juselius, K. (1990) Maximum Likelihood Estimation and Inference on Cointegration—With Applications to the Demand for Money. *Oxford Bulletin of Economics and Statistics*, **52**, 169-210. https://doi.org/10.1111/j.1468-0084.1990.mp52002003.x

[18] Dickey, D.A., Jansen, D.W. and Thornton, D.L. (1994) A Primer on Cointegration with an Application to Money and Income. In: Rao, B.B., Ed., *Cointegration*, Palgrave Macmillan, London, 9-45. https://doi.org/10.1007/978-1-349-23529-2_2

[19] Shibata, R. (1976) Selection of the Order of an Autoregressive Model by Akaike's Information Criterion. *Biometrika*, **63**, 117-126. https://doi.org/10.1093/biomet/63.1.117

[20] Engle, R.F. and Granger, C.W.J. (1987) Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, **55**, 251-276. https://doi.org/10.2307/1913236

[21] Marquez, J. (1995) Time Series Analysis: James D. Hamilton, 1994, (Princeton University Press, Princeton, NJ), 799 pp., US $55.00, ISBN 0-691-04289-6. *International Journal of Forecasting*, **11**, 494-495. https://doi.org/10.1016/0169-2070(95)90035-7

[22] Granger, C.W.J. (1969) Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica*: *Journal of the Econometric Society*, **37**, 424-438. https://doi.org/10.2307/1912791

Appendix

Glossary of Abbreviations

ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criteria
CRAN	Comprehensive R Archive Network
GDP	Gross Domestic Product
IHS	Information Handling Services
ISP	Interpretive Software Products
JMP	John's Macintosh Project
LKR	Sri Lankan Rupee
NIST	National Institute of Standards and Technology
ODBC	Open Database Connectivity
OS	Operating System
PP	Phillips Perron
QMS	Quantitative Micro Software
RATS	Regression Analysis of Time Series
SBIC	Schwarz Bayesian Information Criteria
SPSS	Statistical Package for the Social Sciences
SS	Statistical Software
TSP	Time Series Processor
USD	United States Dollar
VAR	Vector Autoregressive
VECM	Vector Error Correction Model