

# The Contribution of Edgeworth, Bootstrap and Monte Carlo Methods in Financial Data

Edlira Donefski, Tina Donefski, Lorenc Ekonomi

**Abstract**—Edgeworth Approximation, Bootstrap and Monte Carlo Simulations have a considerable impact on the achieving certain results related to different problems taken into study. In our paper, we have treated a financial case related to the effect that have the components of a Cash-Flow of one of the most successful businesses in the world, as the financial activity, operational activity and investing activity to the cash and cash equivalents at the end of the three-months period. To have a better view of this case we have created a Vector Autoregression model, and after that we have generated the impulse responses in the terms of Asymptotic Analysis (Edgeworth Approximation), Monte Carlo Simulations and Residual Bootstrap based on the standard errors of every series created. The generated results consisted of the common tendencies for the three methods applied, that consequently verified the advantage of the three methods in the optimization of the model that contains many variants.

**Keywords**—Autoregression, Bootstrap, Edgeworth Expansion, Monte Carlo Method.

## I. INTRODUCTION

THE econometrics is related to the application of some statistical methods in the solution of an economical problem. The difference between the econometrist's and statistician's points of view has to do with the stochastic relation that is essentially considered by the econometrist. To solve the financial problem of this study related to the influence of Cash-Flow's components in the Cash and Cash Equivalents, we have created a VAR Model (Vector Autoregression Model Estimate) with the series created for the financial activities, operational activities, investing activities and cash and cash equivalent for every three months' end during 30 years, through the EViews10 software package. An abbreviated description of the treatments of the basics of VAR analysis and the important features of the model is given in [9]. Lag Tests and Lag Length Criteria compute various criteria to select the lag order of an unrestricted VAR and it is necessary to specify the maximum lag to "test" for. The results display various information criteria for all lags up to the specified maximum [8]. Lagrange Multiplier test (LM test) statistics for residual serial correlation till the specified order, in the multivariate case is used by the Autocorrelation LM

Test. The version of this LM test is the LR's formula [14, p.22]. Edgeworth expansion correction is applied by the shape of this statistics [4]. The Rao F-test is another version of LM statistics, except the LR version that can be computed in EViews [4]. The simulations of Edgerton and Shukur [4] suggest that it performs best among the many variants they consider. The impulse response is related to a one-time influence to one of the innovations on current and future values of the variables taken in [15], [16]. In the impulse responses, one of the most important commands that gives the better view of the VAR model estimation is the Asymptotic Analysis that is related with the Edgeworth Approximation [1], [5], [10]. Also, Monte Carlo is one of the best methods that improve the view of generated results after the realized replications for limited observations. This method follows the same logic as the Bootstrap method and the both of them contribute to more certain results.

In our paper, after the process of completing the condition of the stationary series, in order to be predictable, we tested the lags to reduce the serial correlation and heteroskedasticity problems. At the end we studied the impulse response of the VAR model related to the standard deviation. In this field we used the asymptotic analysis that means the Edgeworth Approximation Model, Monte Carlo and Bootstrap, whose results presented a common distribution of the responses of all series after the shocks in the standard errors of the other series.

## II. VECTOR AUTOREGRESSIONS

The VAR is widely applied for the prognosis of the systems of interrelated time series and for analyzing the dynamic influence of random shocks on the system of variables. The reduced form VAR is not focused on the structural modeling by treating every endogenous variable of the system as a function of p-lagged values of all of the endogenous variables included in the system [15], [16]. The treatments of the basics of VAR analysis and the important features of the model are described in [9].

We may write the stationary,  $k$ -dimensional,  $\text{VAR}(p)$  process as:

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + Cx_t + \varepsilon_t \quad (1)$$

where  $y_t = (y_{1t}, y_{2t}, \dots, y_{kt})'$  is a  $k \times 1$  vector of endogenous variables,  $x_t = (x_{1t}, x_{2t}, \dots, x_{dt})'$  is a  $d \times 1$  vector of exogenous variables,  $A_1, \dots, A_p$  are  $k \times k$  matrices of lag coefficients to be estimated,  $C$  is a  $k \times d$  matrix of exogenous variable coefficients to be estimated,  $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{kt})'$  is a  $k \times 1$  white noise

E. Donefski is with the Faculty of Natural and Human Sciences, Department of Mathematics and Physics, "Fan S. Noli" University, Korça, Albania (e-mail: edonefski@unkorce.edu.al)

T. Donefski is with the Faculty of Economy, Department of Finance-Accounting, "Fan S. Noli" University, Korça, Albania (e-mail: tdonefski@unkorce.edu.al)

L. Ekonomi, is with the Faculty of Natural and Human Sciences, Department of Mathematics-Informatics-Physics, "Fan S. Noli" University, Korça, Albania (e-mail: lorencekonomi@yahoo.co.uk)

innovation process, with,

$$E(\varepsilon_t) = 0, E(\varepsilon_t \varepsilon_s') = \Sigma_\varepsilon, \text{ and } E(\varepsilon_t \varepsilon_s') = 0, \text{ for } t \neq s.$$

This equation shows that the vector of innovations is conjointly correlated with full rank matrix, but not correlated with their leads and lags of the innovations and the right-hand side variables [15], [16].

Null Hypothesis: D(CASH&CASH EQUIVALENTS) has a unit root  
Exogenous: Constant  
Lag Length: 2 (Automatic - based on SIC, maxlag=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-10.27355	0.0000
Test critical values:		
1% level	-3.490772	
5% level	-2.887909	
10% level	-2.580908	

\*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation  
Dependent Variable: D(CASH&CASH EQUIVALENTS,2)  
Method: Least Squares  
Date: 02/14/21 Time: 13:43  
Sample (adjusted): 5 114  
Included observations: 110 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(CASH&CASH EQUIVALENTS(-1))	-2.482254	0.241616	-10.27355	0.0000
D(CASH&CASH EQUIVALENTS(-1),2)	0.751484	0.180939	4.153233	0.0001
D(CASH&CASH EQUIVALENTS(-2),2)	0.294790	0.095085	3.100280	0.0025
C	194.9927	244.3143	0.798122	0.4266
R-squared	0.787872	Mean dependent var	-43.17273	
Adjusted R-squared	0.781868	S.D. dependent var	5461.514	
S.E. of regression	2550.779	Akaike info criterion	18.56187	
Sum squared resid	6.90E+08	Schwarz criterion	18.66007	
Log likelihood	-1016.903	Hannan-Quinn criter.	18.60170	
F-statistic	131.2324	Durbin-Watson stat	2.056134	
Prob(F-statistic)	0.000000			

Fig. 1 Stationary of the series taken into study

Let the  $(pk+d) \times 1$  vector,  $Z_t = (y_{t-1}', \dots, y_{t-p}', x_t')$ , show all the  $t$  period regressors in the VAR model. Also, for observations, this model can be written as:  $Y = BZ + E$  where  $Y = (y_1, y_2, \dots, y_T)$  and  $E = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T)$  are  $k \times T$  matrices of endogenous variables and innovations, and,

$$B = (A_1, A_2, \dots, A_p, C)$$

$$Z = (Z_1, Z_2, \dots, Z_T)$$

are the  $k(pk+d)$  matrix of system coefficients and the  $(pk+d) \times T$  matrix of regressor data, respectively. In stacked form, we have  $\tilde{y} = (Z' \otimes I_k) \beta + \tilde{\varepsilon}$  where  $\tilde{y} = \text{vec}(Y)$ ,  $\beta = \text{vec}(B)$ , and

$$\tilde{\varepsilon} = \text{vec}(E), \text{ where } E(\tilde{\varepsilon} \tilde{\varepsilon}') = (I_k \otimes \Sigma_\varepsilon).$$

As long as the lagged values of the endogenous variables and the innovations are supposed not to be correlated with lagged innovations and the exogenous variables, the standard conditions are supported and the OLS curve gives a stable evaluation.

The efficiency of the OLS model that basically is equal to GLS, is supported by the fact that the all of the equations in the system have identical regressors.

In Eviews, the simple Ordinary Least Square Estimation is applied to each equation, for generating the assessment of the standard VAR model [15], [16]. Implementing OLS estimation to the shelf yields the LS estimator  $\hat{\beta} = ((ZZ')^{-1} Z' \otimes I_k) \tilde{y}$  which has covariance matrix  $V(\hat{\beta}) = (ZZ')^{-1} \otimes \Sigma_\varepsilon$ .

To get a covariance matrix of evaluation we need an assessment, which is usually obtained by using d.f. corrected estimator of the remaining (residual) moment:

$$\hat{\Sigma}_\varepsilon = \frac{\hat{E} \hat{E}'}{T - (pk + d)}$$

where  $\hat{E} = Y - \hat{B}Z$  for  $\hat{\beta} = \text{vec}(\hat{B})$ .

### EViews 10

EViews is a statistical and econometric software package. The most current professional version is EViews 10, because it helps in creating the aiming model and studying it through a various commands and scripts. We have done the study of our model with the help of this software.

### III. LAG TESTS

#### Lag Exclusion Tests

The Lag Exclusion Tests perform the default exception tests for every lag of the VAR model. The (Wald) statistic gives a separated and common report of each equation related to the joint importance of all endogenous variables at every one of the lags included in the model.

#### Lag Length Criteria

The Lag Length Criteria generate different criteria of the lag order selection for an unrestricted VAR. It is necessary to specify the maximum order of lags for realizing the test. The generated results show different information criteria for all lags for the determined maximum. (The lag starts at 1 if the VAR model is free of exogenous variables and it starts at 0 for the inverse alternative) The criteria are explained in more details in [8]. The LR test begins with the maximum lag and uses the statistics to test the hypothesis that states for the commonly zero value of the coefficients on lag  $l$ :

$$LR = (T - m) \{ \log | \Sigma_{\varepsilon, l-1} | - \log | \Sigma_{\varepsilon, l} | \} \sim \chi^2(k^2)$$

where the number of parameters is represented by  $m$ . The Sims' [3] sample modification which uses  $(T-m)$  instead of  $T$  can be considered too. The modified LR statistics result should be compared to the 5% critical value that starts from the maximum lag, and decreases the lag one at a time until getting the first rejection. If the tests do not identify rejects, the minimum lag will be noted with an asterisk. An important conclusion is that even though the individual tests have size 0.05, the whole size of the test is impossible to be 5%; see the discussion in [8, p.125-126], [15], [16].

(d)

39

Sample: 1 114  
Included observations: 107

Null hypothesis: No serial correlation at lag h

Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	19.71434	16	0.2334	1.246910	(16, 229.8)	0.2337
2	17.07720	16	0.3806	1.074038	(16, 229.8)	0.3810
3	22.78647	16	0.1196	1.450740	(16, 229.8)	0.1198
4	32.42915	16	0.0088	2.107965	(16, 229.8)	0.0888
5	21.64000	16	0.1552	1.374364	(16, 229.8)	0.1555
6	19.52886	16	0.2422	1.234688	(16, 229.8)	0.2426
7	13.28754	16	0.6516	0.828962	(16, 229.8)	0.6519

Null hypothesis: No serial correlation at lags 1 to h

Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	19.71434	16	0.2334	1.246910	(16, 229.8)	0.2337
2	38.33890	32	0.2040	1.215297	(32, 263.4)	0.2054
3	52.56696	48	0.3016	1.105646	(48, 260.1)	0.3060
4	75.42465	64	0.1554	1.202771	(64, 248.9)	0.1622
5	100.0980	80	0.0637	1.295347	(80, 235.2)	0.0707
6	109.2063	96	0.1685	1.158040	(96, 220.4)	0.1901
7	112.1851	112	0.4773	0.987564	(112, 205.1)	0.5235

\*Edgeworth expansion corrected likelihood ratio statistic.

Fig. 3 VAR residual serial correlation LM tests

VAR Residual Heteroskedasticity Tests (Levels and Squares)

Date: 02/14/21 Time: 13:54

Sample: 1 114

Included observations: 107

Joint test:

Chi-sq	Df	Prob.
573.0495	480	0.2022

Individual components:

Dependent	R-squared	F(48,58)	Prob.	Chi-sq(48)	Prob.
res1*res1	0.226286	0.353398	0.9998	24.21260	0.9984
res2*res2	0.374411	0.723181	0.8753	40.06201	0.7854
res3*res3	0.317155	0.561224	0.9794	33.93559	0.9377
res4*res4	0.342811	0.630305	0.9492	36.68075	0.8833
res2*res1	0.285315	0.482389	0.9948	30.52874	0.9768
res3*res1	0.219804	0.340423	0.9999	23.51903	0.9989
res3*res2	0.354825	0.664543	0.9266	37.96625	0.8500
res4*res1	0.226309	0.353445	0.9998	24.21508	0.9983
res4*res2	0.365132	0.694950	0.9020	39.06913	0.8175
res4*res3	0.252436	0.408028	0.9991	27.01067	0.9938

Fig. 4 VAR residual heteroscedasticity tests

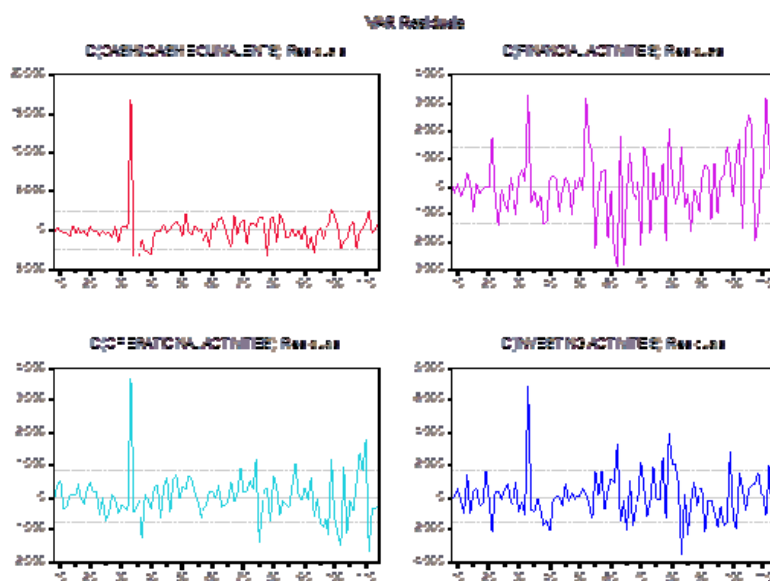


Fig. 5 VAR residuals

#### IV. RESIDUAL TESTS

##### Autocorrelation LM Test

LM test statistics for residual serial correlation till the specified order in the multivariate case is used by the Autocorrelation LM Test. The statistical test of Breusch-Godfrey, which is used in identifying the autocorrelation at lag order  $h$ , can be computed by executing an adjuvant regression of the residuals  $u_t$  on the original right-hand regressors and the lagged residual  $u_{t-h}$  and the first  $h$  values of  $u_{t-h}$  that are missed are stuffed with zeros.

The version of this LM test is the LR's formula [14, p.22]. Edgeworth expansion correction is applied by the shape of this statistics [4]. The LM statistic is asymptotically distributed  $\chi^2$  with  $k^2$  freedom's degree, in the case of no serial correlation of order  $h$ , as a null hypothesis.

The Rao F-test is another version of LM statistics, except

the LR version that can be computed in EViews [4]. The simulations of Edgerton and Shukur [4] suggest that it performs best than the many other variants they consider.

##### White Heteroskedasticity Test

The test regression is gained through testing the joint significance of the regression after regressing every cross generation of the residuals on the cross products of the regressors. The No Cross Terms alternative is based only on the levels and squares of the original regressors, whereas the With Cross Terms option involves all non-redundant cross-generations of the original regressors in the test equation. A constant term can be included as a regressor in a regression test.

The first results show the common importance of the regressors with the exception of the constant term. These tests verify the stability of every separated element in the residual covariance matrix. The null hypothesis shows the assumption

that the regressors that are not constant have not a common significance.

Response of D(CASH&CASH EQUIVALENTS):				
Period	D(CASH&CASH EQUIVALENTS)	D(FINANCIAL ACTIVITIES)	D(INVESTING ACTIVITIES)	D(OPERATIONAL ACTIVITIES)
1	2394.841 (163.708)	0.000000 (0.000000)	0.000000 (0.000000)	0.000000 (0.000000)
2	-1826.573 (289.309)	-157.7709 (265.879)	-10.38674 (256.358)	-411.8117 (261.747)
3	134.3653 (328.880)	153.8764 (326.177)	-87.20782 (327.976)	165.1840 (315.895)
4	-29.87395 (329.661)	-466.9623 (321.799)	330.6794 (305.294)	-179.7698 (322.856)
5	-233.3102 (335.568)	-314.7734 (303.122)	-76.14772 (184.436)	399.0172 (331.251)
6	297.3177 (339.679)	554.6333 (307.600)	-183.4638 (197.308)	-191.1481 (337.646)
7	49.41530 (338.443)	74.94726 (309.923)	-28.38113 (195.673)	161.8234 (346.212)
8	135.6295 (336.174)	82.87116 (296.047)	144.2722 (181.838)	141.3084 (275.995)
9	-284.6496 (237.246)	-78.07673 (215.872)	-41.52852 (146.351)	-349.2871 (245.257)
10	-62.83569 (183.436)	-166.1535 (227.294)	60.96567 (145.817)	197.3186 (221.519)
Response of D(FINANCIAL ACTIVITIES):				
Period	D(CASH&CASH EQUIVALENTS)	D(FINANCIAL ACTIVITIES)	D(INVESTING ACTIVITIES)	D(OPERATIONAL ACTIVITIES)
1	452.1291 (128.311)	1288.179 (88.0581)	0.000000 (0.000000)	0.000000 (0.000000)
2	-351.0695 (158.861)	-537.1552 (155.530)	183.6154 (145.278)	-106.5981 (148.527)
3	-130.6953 (170.712)	-507.9599 (165.972)	23.61292 (164.415)	-50.76067 (160.764)
4	92.96241 (172.659)	-79.68856 (169.289)	-118.5471 (152.975)	37.34091 (185.526)
5	-4.760887 (181.465)	511.8143 (172.489)	74.25473 (96.4010)	153.5016 (185.797)
6	217.8045 (183.054)	-195.4912 (174.402)	4.729393 (111.332)	-54.24950 (207.385)
7	-62.22006 (182.373)	51.91789 (178.470)	-37.03803 (109.356)	32.10449 (209.134)
8	-97.76470 (179.007)	41.81660 (177.297)	-182.4105 (89.1186)	38.38268 (163.269)
9	-102.6911 (113.010)	37.31912 (157.644)	282.0539 (87.5493)	-121.5019 (159.113)
10	74.33718 (108.801)	-339.0421 (160.595)	17.39194 (88.5959)	-1.790863 (121.752)

Response of D(OPERATIONAL ACTIVITIES):				
Period	D(CASH&CASH EQUIVALENTS)	D(FINANCIAL ACTIVITIES)	D(INVESTING ACTIVITIES)	D(OPERATIONAL ACTIVITIES)
1	344.1629 (72.2384)	-98.12310 (67.9698)	699.6536 (47.8273)	0.000000 (0.000000)
2	-315.0287 (92.8589)	-22.40695 (91.9093)	-329.7068 (86.1053)	-0.254375 (85.3943)
3	16.49977 (100.606)	112.8611 (99.6693)	-169.4401 (99.7639)	153.0970 (95.8599)
4	51.02785 (102.746)	-96.27422 (100.761)	-196.2446 (92.3634)	-46.59434 (104.871)
5	-75.44071 (119.201)	-5.808033 (113.973)	608.0812 (75.4234)	-2.434657 (107.397)
6	87.91146 (127.496)	61.75002 (122.946)	-370.4441 (94.3426)	-65.27948 (132.626)
7	-48.70167 (129.149)	-6.001829 (126.140)	-102.4900 (97.9135)	117.3536 (136.154)
8	32.71189 (128.583)	-96.41442 (125.259)	-177.7790 (75.3909)	-35.30156 (116.996)
9	-49.14106 (112.800)	114.8115 (130.570)	637.3146 (82.4227)	-146.5617 (112.711)
10	61.76571 (114.588)	-57.35303 (138.660)	-323.5077 (101.816)	41.83503 (132.626)

Response of D(INVESTING ACTIVITIES):				
Period	D(CASH&CASH EQUIVALENTS)	D(FINANCIAL ACTIVITIES)	D(INVESTING ACTIVITIES)	D(OPERATIONAL ACTIVITIES)
1	721.3286 (150.392)	-226.4929 (141.233)	470.3973 (136.649)	1373.819 (93.9123)
2	-832.6242 (189.381)	340.2930 (182.608)	-235.0604 (174.004)	-71.17374 (177.963)
3	42.87032 (206.734)	306.5980 (205.562)	155.5651 (205.748)	-341.1764 (202.845)
4	-24.01338 (212.788)	-499.4195 (209.133)	-107.3987 (192.826)	-278.4853 (217.033)
5	-234.5903 (218.630)	-265.0160 (211.758)	83.12071 (145.528)	-164.3941 (226.157)
6	198.1058 (221.606)	-22.04405 (211.201)	-311.1000 (148.224)	287.5668 (234.811)
7	241.8573 (221.090)	186.2475 (209.659)	-139.6171 (138.017)	-26.57571 (243.331)
8	236.7618 (220.739)	291.3367 (208.306)	8.846908 (139.841)	-120.1115 (207.571)
9	-220.7173 (159.639)	43.25884 (179.017)	312.2315 (118.895)	-2.796193 (203.625)
10	-210.8424 (145.215)	-225.1515 (175.564)	-54.15937 (119.570)	-33.54819 (184.544)

Cholesky Ordering: D(CASH&CASH EQUIVALENTS)  
D(FINANCIAL ACTIVITIES) D(OPERATIONAL ACTIVITIES)  
D(INVESTING ACTIVITIES)

Fig. 6 Asymptotic analysis (Edgeworth Approximation) for the identification of the particular series response

## V. IMPULSE RESPONSES

The impulse responses explain the effect of an immediate change to the  $i$ -th variable, that influences the other variables and this effect is transmitted to all other endogenous variables included in the VAR model. The impulse response is related to a one-time influence to one of the innovations on current and future values of the variables taken in [15], [16].

The result of the impulse response is clear, if the innovations are conjointly not correlated. The  $i$ -th innovation is just an immediate change to the  $i$ -th endogenous variable. Innovations are usually correlated, and cannot be associated with a specific variable [15], [16]. For a right interpretation of the impulses, it is necessary to implement a transformation to the innovations for making them uncorrelated:  $u_i = P\varepsilon_i \sim (0, D)$  where  $D$  is a diagonal covariance matrix. EVIEWS provides several options for the choice of  $P$  which produce a diagonal matrix.

In the impulse responses, one of the most important commands that give the better view of the VAR model estimation is the Asymptotic Analysis that is related with the Edgeworth Approximation.

## Edgeworth Expansion

Let  $X_1, X_2, \dots, X_n$  be independent and identically distributed random variables with mean  $\mu$  and variance  $\sigma^2$ . By the Central Limit Theorem,

$$S_n = \frac{\sum_{i=1}^n X_i / n - \mu}{\sigma / \sqrt{n}}$$

is asymptotically normally distributed with zero mean and unit variance. We are interested in the asymptotic behavior of the difference between the normal distribution  $\Phi(x)$  and the distribution function  $F_n(x)$  of the  $S_n$ . By logarithmic expansion, characteristics functions, using the expansion series of the exponential function and doing the necessary transformations [2], [6], [11], [13], we can get:

$$P(S_n \leq x) = \Phi(x) + n^{-1/2} p_1(x) \phi(x) + n^{-1} p_2(x) \phi(x) + \dots + n^{-j/2} p_j(x) \phi(x) + \dots \quad (2)$$

called the Edgeworth expansion of the distribution of

$P(S_n \leq x)$ . Here  $\Phi$  denotes the standard normal distribution function,  $\phi$  denotes the standard normal density. The polynomial  $p_j$  has degree of order  $3j-1$  and is odd for even  $j$ . Hence,

$$p_1(x) = -\frac{1}{6}k_3(x^2 - 1)$$

and,

$$p_2(x) = -x \left\{ \frac{1}{24}k_4(x^2 - 3) + \frac{1}{72}k_3^2(x^4 - 10x^2 + 15) \right\}.$$

The third cumulant  $k_3$  refers to skewness, so the term of  $n^{-1/2}$  order improves the basic normal approximation of the cumulative distribution function of  $S_n$  by performing skewness correction.  $k_4$  refers to kurtosis for the term of order  $n^{-1}$  which improves the normal approximation further by adjusting for kurtosis.

Usually (2) exists as an asymptotic series, which means that if the series stop at a specific order the remainder is of smaller order than the last omitted term in the series. It means:

$$P(S_n \leq x) = \Phi(x) + n^{-1/2}p_1(x)\phi(x) + n^{-1}p_2(x)\phi(x) + \dots + n^{-j/2}p_j(x)\phi(x) + o(n^{-j/2}) \quad (3)$$

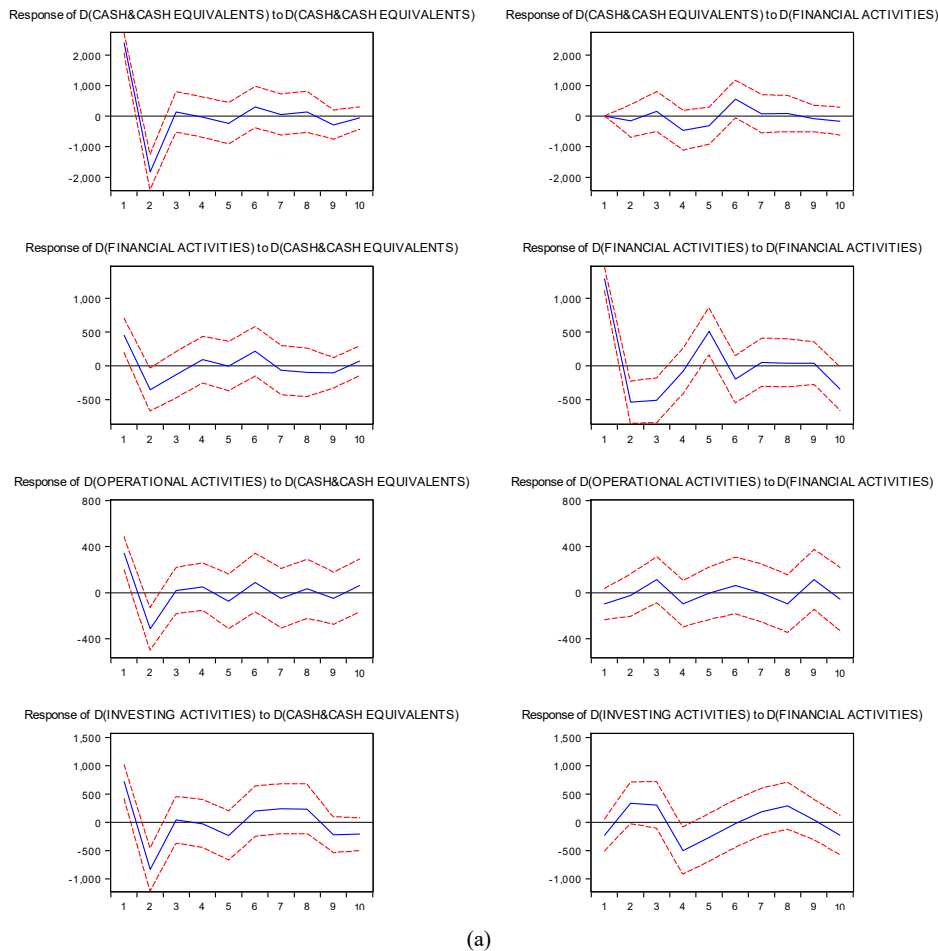
The restrictions on (2) are:

$$E(|X|^{j+2}) < \infty \text{ and } \lim_{t \rightarrow \infty} \sup |\psi(t)| < 1.$$

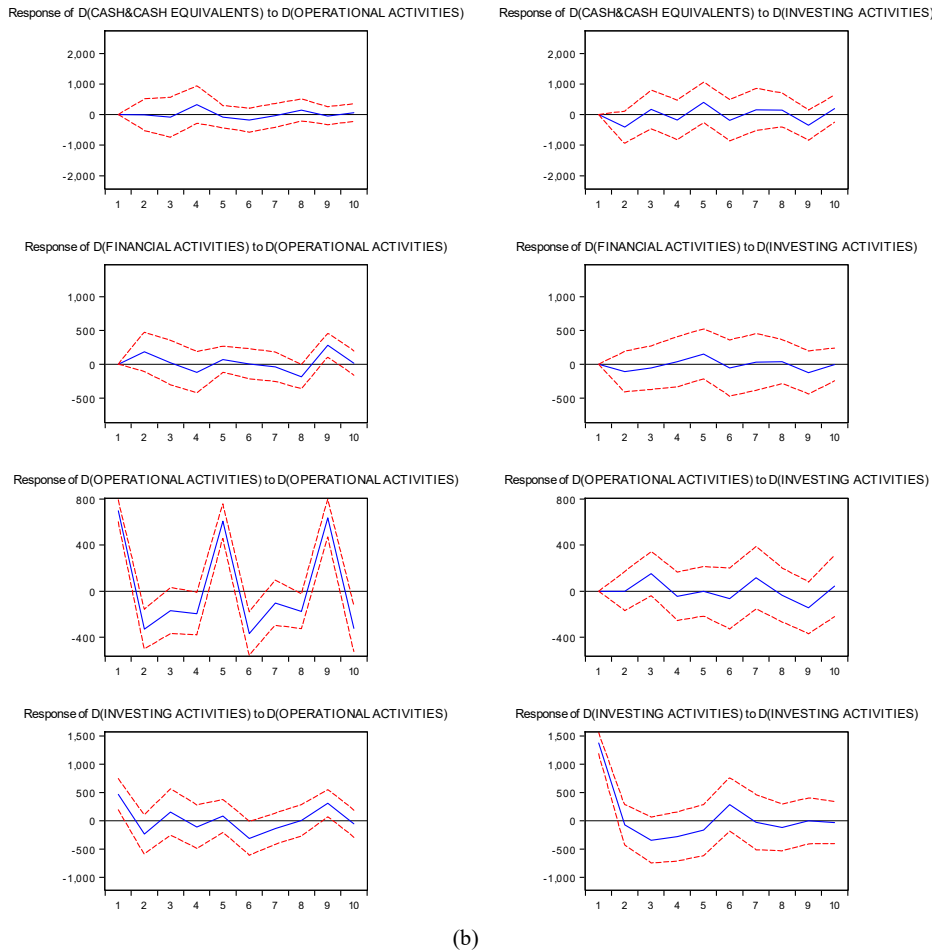
One can find the proof of this fact in [10].

The techniques used to derive Edgeworth expansions for the distribution of one-dimensional statistics can be generalized to the multivariate case. However, in the multivariate case the notation becomes more complex. A basic result on multivariate Edgeworth expansions is that of [12].

Now, let see the order of Edgeworth expansions for a stationary autoregressive series.



(a)



(b)

Fig. 7 The graphical view of the asymptotic analysis of the response of the particular series

Let  $Y_t$  be a stationary autoregressive process satisfying,

$$Y_t = \sum_{i=1}^p \theta_i Y_{t-i} + \varepsilon_t,$$

where we assume that:

- (A.1)  $(\varepsilon_t)$  are i.i.d.  $\sim F_0, E\varepsilon_t = 0, E\varepsilon_t^2 = 1, E\varepsilon_t^{2(s+1)} < \infty$  for some  $s \geq 3$ .
- (A.2)  $(\varepsilon_1, \varepsilon_1^2)$  satisfies Cramer's condition, i.e., for every  $d > 0$ , there exists  $\delta > 0$  such that  $\sup_{\|t\| > d} |E \exp(it'(\varepsilon_1, \varepsilon_1^2))| \leq \exp(-\delta)$ .
- (A.3) Roots of  $\sum_{j=0}^p \theta_j z^{p-j} = 0$  lie within the unit circle.

Here  $\theta_0 = 1$ .

The arguments [1], [5] show that if conditions (A.1) and (A.3) hold and  $\varepsilon_1$  satisfies Cramer's condition, then the distribution of  $n^{-1/2} \tilde{S}_n(\theta_n - \theta)$  admits an Edgeworth expansion of order  $o(n^{-(s-2)/2})$ .

#### Monte Carlo Method

Monte Carlo is one of the best methods that improves the view of generated results after the realized replications.

Monte Carlo and Bootstrap versions have common results in the practical data, where they can be applied. The data used in practical studies are limited and this condition is one of the most important ones, because the repeating samples for the particular limited observations gives certain results.

#### Impulse Response Standard Errors and Confidence Intervals

The asymptotic approximation gives a better view of the confidence intervals for a VAR model and this kind of distribution has an optimal performance in limited samples [7].

The impulse response standard errors through generating the Confidence Intervals are related to the application of the Monte Carlo and Bootstrap methods because of the limited observations used. This application gives optimal solution for a case taken in study.



Response of D(CASH&CASH EQUIVALENTS):					Response of D(Operational Activities):				
Period	D(CASH&CASH EQUIVALENTS)	D(FINANCIAL ACTIVITIES)	D(INVESTING ACTIVITIES)	D(OPERATIONAL ACTIVITIES)	Period	D(CASH&CASH EQUIVALENTS)	D(FINANCIAL ACTIVITIES)	D(INVESTING ACTIVITIES)	D(OPERATIONAL ACTIVITIES)
1	2394.841 (170.596)	0.000000 (0.00000)	0.000000 (0.00000)	0.000000 (0.00000)	1	344.1629 (71.5141)	-98.12310 (69.0335)	699.6536 (47.3397)	0.000000 (0.00000)
2	598.2679 (299.914)	-157.7709 (275.133)	-10.38674 (275.943)	-411.8117 (271.467)	2	29.13426 (98.5558)	-120.5300 (96.7324)	369.9468 (89.5173)	-0.254375 (89.0068)
3	702.6332 (300.923)	-3.894493 (304.405)	-97.59456 (285.315)	-246.6277 (300.308)	3	45.63403 (97.1030)	-7.868934 (103.718)	200.5067 (101.969)	152.8427 (105.068)
4	672.7592 (310.196)	-470.8568 (322.099)	233.0849 (302.059)	-426.3975 (307.083)	4	96.66188 (102.578)	-103.9431 (113.316)	4.262022 (109.148)	106.2483 (110.518)
5	439.4490 (329.784)	-785.6302 (349.972)	156.9372 (272.865)	-27.38035 (332.738)	5	21.22117 (125.202)	-109.7512 (135.273)	612.3433 (116.433)	103.8137 (108.598)
6	736.7667 (350.220)	-230.9969 (354.627)	-26.52665 (271.063)	-218.5284 (335.095)	6	109.1326 (133.020)	-48.00116 (141.208)	241.8991 (114.388)	38.53418 (130.332)
7	786.1820 (334.468)	-156.0496 (368.168)	-54.90778 (241.888)	-56.70502 (364.191)	7	60.43096 (116.985)	-54.00299 (133.104)	139.4091 (101.312)	155.8878 (145.947)
8	921.8115 (285.626)	-73.17848 (296.676)	89.36438 (226.182)	84.60340 (296.107)	8	93.14285 (98.3919)	-150.4174 (122.388)	-38.36987 (100.969)	120.5862 (121.972)
9	637.1619 (291.697)	-151.2552 (323.543)	47.83586 (231.283)	-264.6837 (268.044)	9	44.00179 (114.155)	-35.60591 (136.234)	598.9447 (104.093)	-25.97544 (107.473)
10	574.3262 (273.179)	-317.4087 (332.872)	108.9015 (229.498)	-67.36506 (277.699)	10	105.7675 (120.855)	-92.95895 (143.787)	275.4370 (109.605)	15.85959 (120.295)

Response of D(Financial Activities):					Response of D(Investing Activities):				
Period	D(CASH&CASH EQUIVALENTS)	D(FINANCIAL ACTIVITIES)	D(INVESTING ACTIVITIES)	D(OPERATIONAL ACTIVITIES)	Period	D(CASH&CASH EQUIVALENTS)	D(FINANCIAL ACTIVITIES)	D(INVESTING ACTIVITIES)	D(OPERATIONAL ACTIVITIES)
1	452.1291 (134.964)	1288.179 (87.7315)	0.000000 (0.00000)	0.000000 (0.00000)	1	721.3286 (147.927)	-226.4929 (148.659)	470.3973 (141.400)	1373.819 (96.2577)
2	101.0596 (175.728)	751.0235 (170.563)	183.6154 (153.147)	-106.5981 (153.740)	2	-111.2956 (224.226)	113.8001 (235.185)	235.3369 (225.861)	1302.645 (205.996)
3	-29.63566 (184.872)	243.0635 (194.814)	207.2283 (181.139)	-157.3587 (197.890)	3	-68.42523 (254.162)	420.3981 (264.313)	390.9020 (260.061)	961.4687 (264.973)
4	63.32674 (189.137)	163.3950 (195.861)	88.69121 (187.774)	-120.0178 (198.129)	4	-92.43861 (273.604)	-79.02140 (284.494)	283.5033 (290.718)	682.9834 (286.524)
5	58.56586 (205.685)	675.2093 (218.795)	162.9359 (180.737)	33.48376 (200.104)	5	-327.0289 (290.705)	-344.0374 (324.814)	366.6240 (301.664)	518.5893 (302.668)
6	276.3703 (203.677)	479.7182 (239.033)	167.6653 (166.719)	-20.76574 (213.051)	6	-128.9231 (298.297)	-366.0815 (337.033)	55.52404 (305.522)	806.1561 (316.404)
7	214.1503 (198.085)	531.6361 (227.604)	130.6273 (141.865)	11.33875 (235.374)	7	112.9342 (284.857)	-179.8340 (355.210)	-84.09303 (276.658)	779.5804 (348.108)
8	116.3856 (166.907)	573.4527 (211.207)	-51.76321 (151.198)	49.72144 (206.345)	8	349.6960 (224.383)	111.5027 (307.111)	-75.24612 (244.770)	659.4689 (331.971)
9	13.69447 (178.023)	610.7718 (223.638)	230.2707 (153.458)	-71.78050 (173.802)	9	128.9787 (221.602)	154.7615 (293.891)	236.9854 (246.248)	656.6727 (291.401)
10	88.03166 (181.130)	271.7297 (240.313)	247.6626 (170.350)	-73.57136 (180.248)	10	-81.86372 (237.247)	-70.38999 (299.921)	182.8260 (239.909)	623.1245 (264.664)

Cholesky Ordering: D(CASH CASH EQUIVALENTS)

(a)

(b)

Fig. 8 Monte Carlo results, for 1000 replications for the identification of the particular series response

## VI. THE VAR BOOTSTRAP

If a system of time series can be modified as a VAR model of order  $p$ , the process of data generation for this system can be expressed as:

$$y_t = v + \sum_{j=1}^p A_j y_{t-j} + u_t \quad (4)$$

where the  $K$ -element vectors as  $y_t$  and  $v$ ,  $A_1, \dots, A_p$  can be presented as  $K \times K$  matrices form; and  $K \times K$  covariance matrix as  $\Sigma_u$  is taken from  $u_t \sim N(0, \Sigma_u)$ .

The next assumption has to do with the functionality of the aimed object like: estimator, statistical tests etc.,

$$q_T = q(y_1, \dots, y_T, v, A_1, \dots, A_p) \quad (5)$$

The distribution  $F_{q,T}$  of  $q_T$  may be approximated by simulating a considered number, for example  $B$ , of different results ( $q_T^{*b}$  for  $b = 1, \dots, B$ ), where the “\*” is used to show the derivate of the object after the simulations. Also, the  $B$  outputs  $q_T^{*b}$  are taken by the simulations from the process of data generations in (4) to give  $B$  independent results  $y_t^{*b}$  of  $y_t$  and for each output, find the corresponding  $q_T^{*b}$  using

(5) but with the generated data. The distribution of the  $q_T$  is similar as the empirical distribution of the simulated outputs  $\{y_t^{*b}\}_{b=1}^B$ .

## VII. SIMULATIONS OF EDGEWORTH, BOOTSTRAP AND MONTE CARLO METHODS IN FINANCIAL DATA

In this paper, we are going to discuss the relation between the cash & cash equivalents at the end of every three-months during the 30 years, and the financial, operational and investing activities. This information is taken from the financial reported information of the one of the most successful businesses in the world: Coca-Cola. At first, we have tested the stationarity of the series taken into study, as it can be easily identified in Fig. 1. If the condition of stationarity is fulfilled, we are able to study correctly the series of the model.

After the application of the Unit Root Test, it can be easily identified that the differenced form of the Cash & Cash Equivalents series that we have considered in this study is stationary, because the probability of the Augmented Dickey-Fuller test statistics is 0.0000 that is lower than the criteria of 0.05. So, in the VAR model we are going to use the differenced data.

In the graphical view of Figs. 2 (a)-(d) are shown the results of every series' importance in the other's series, where these

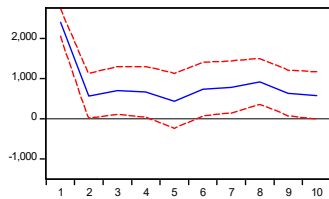


influences are explained by three rows: the first one expresses the coefficient of the two series, the second one is related to the standard error of the relation between the two series and the third one is the t-statistics that is gained through dividing the coefficient with the particular standard errors. For example, the differenced form of Cash & Cash Equivalents for 1 lag has limited importance in the D(Financial Activities), D(Operational Activities) and D(Investing Activities). Even though when the D(Cash&Cash Equivalents)(-1) increases by 1%, the D(Financial Activities), D(Investing Activities) and D(Operational Activities) will decrease particularly by 9.2%, 5.4%, 33.3%. The same logic can be used to explain other relations. The four series that are considered in this study have a considered percentage in the explanation of the VAR

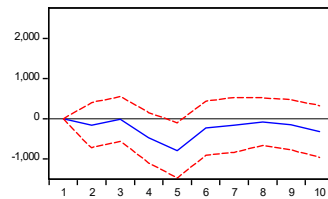
estimated model. So, the D(Cash&Cash Equivalents) can explain the model at 56.8%, the D(Financial Activities) at 75.75%, D(Investing Activities) at 96.69% and D(Operational Activities) at 61.16%.

To further study the model, it is necessary to identify whether the model is free of heteroskedasticity and autocorrelation. So, in Fig. 3 is presented the VAR Residual Correlation through the LM Test that verifies that the model is free of serial correlation at lag  $h$ , because all the probabilities at the lags from 1 to 7 that we took in the study are greater than 0.05. In this way, the Null Hypothesis will not be accepted. Even in the simulations of Edgeworth that perform better among the many alternatives they consider, the result is the same: The estimated VAR Model is free of Correlation.

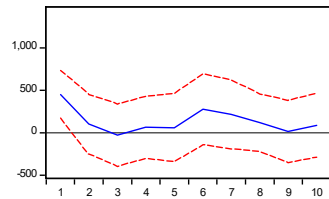
Accumulated Response of D(CASH&amp;CASH EQUIVALENTS) to D(CASH&amp;CASH EQUIVALENTS)



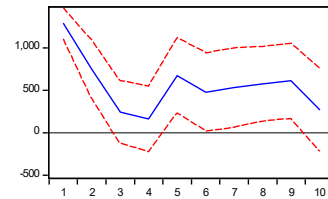
Accumulated Response of D(CASH&amp;CASH EQUIVALENTS) to D(FINANCIAL ACTIVITIES)



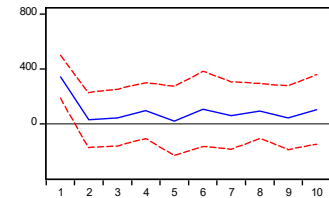
Accumulated Response of D(FINANCIAL ACTIVITIES) to D(CASH&amp;CASH EQUIVALENTS)



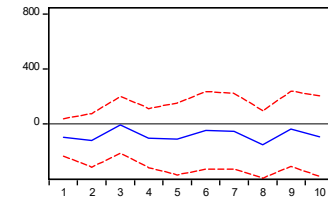
Accumulated Response of D(FINANCIAL ACTIVITIES) to D(FINANCIAL ACTIVITIES)



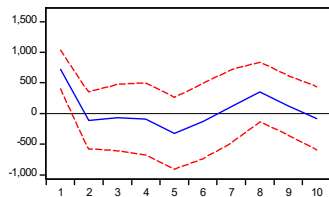
Accumulated Response of D(OPERATIONAL ACTIVITIES) to D(CASH&amp;CASH EQUIVALENTS)



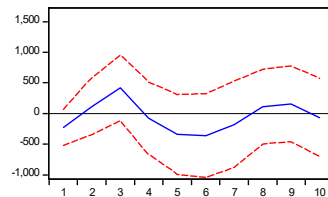
Accumulated Response of D(OPERATIONAL ACTIVITIES) to D(FINANCIAL ACTIVITIES)



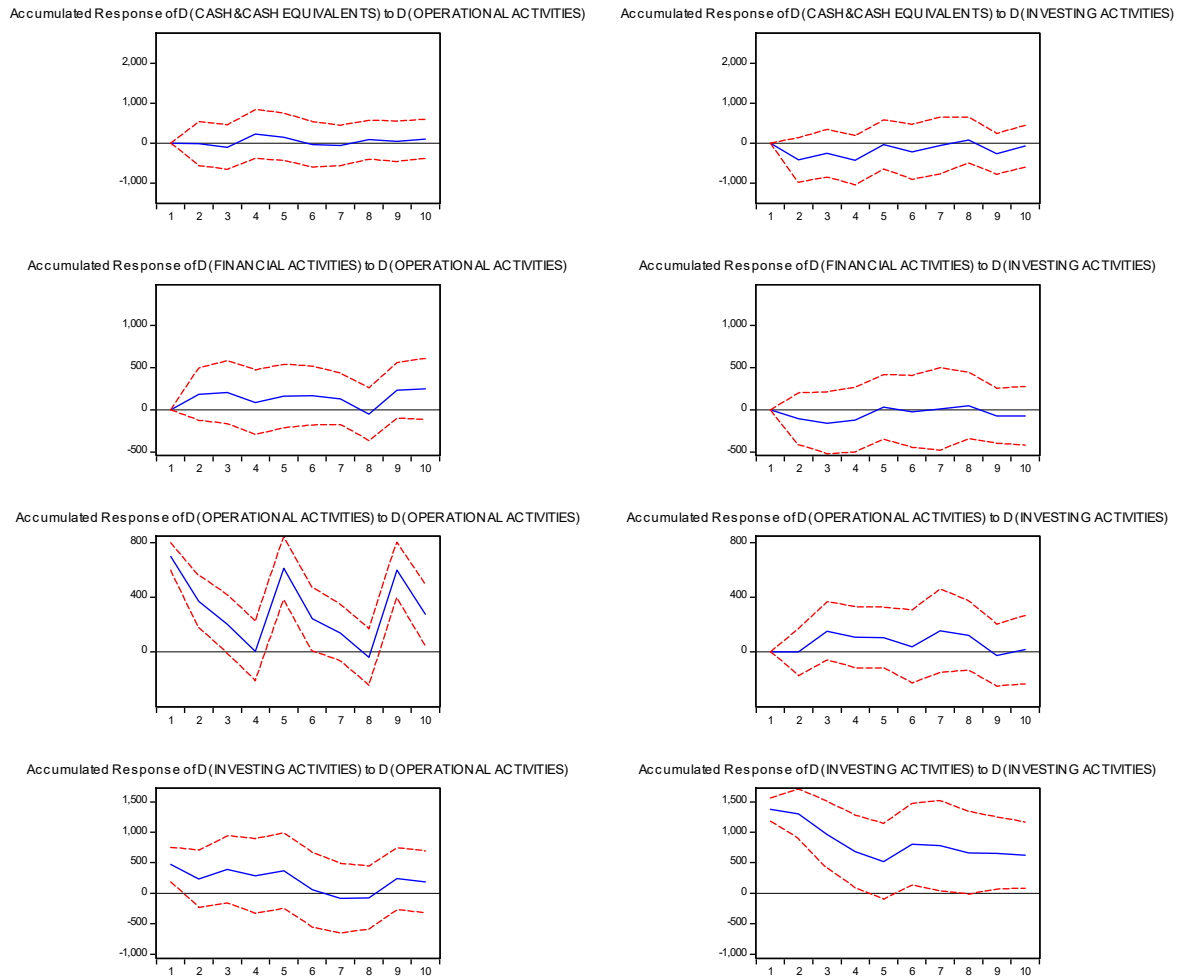
Accumulated Response of D(INVESTING ACTIVITIES) to D(CASH&amp;CASH EQUIVALENTS)



Accumulated Response of D(INVESTING ACTIVITIES) to D(FINANCIAL ACTIVITIES)



(a)



(b)

Fig. 9 The graphical view of the results treated in Figs. 8 (a) and (b)

In Fig. 4, it is easy to identify that the probabilities of every residuals' dependence study are greater than 0.05, and this shows that the null hypothesis expressing the existence of heteroskedasticity cannot be accepted. Thus, the proposed model has homoscedasticity.

To make the estimated VAR model easier to be understood, we have presented the Impulse Responses, where an important issue was the asymptotic analysis that means the Edgeworth Approximation Model, Monte Carlo Approximation for 1000 replications and Bootstrap Version for 1000 Replications and for 95% Confidence.

In Figs. 6 (a) and (b) are presented the asymptotic analyses (Edgeworth Approximation) for the identification of the particular series response, after an increase in standard error of the other influencing series. For each period chosen (10 periods) are presented particular standard error for particular series under the particular series' response.

Figs. 7 (a) and (b) show the graphical views of the asymptotic analysis of the response of the particular series based on the increases of one Standard Error (Deviation) in

other particular, influencing series. So, an increase in Financial Activities by one standard error will trigger a rise in Cash & Cash Equivalents during the fourth quarter. On the other side, an increase in Operational Activities by one standard error will cause a decrease in the Cash & Cash Equivalent series. The same logic can be followed to interpret all the other cases.

In Figs. 8 (a) and (b) are presented the Monte Carlo results, for 1000 replications for the identification of the particular series response, after an increase in standard error of the other influencing series. For each period chosen (10 periods) are presented standard errors for particular series under the particular series' response. It is identified that the results have not a considerable difference compared with the asymptotic analysis (Edgeworth Approximation), however, the small difference is because of the improved results that are given from Monte Carlo simulations that use big number of replications to generate the results.

Figs. 9 (a) and (b) present the graphical view of the results treated in Figs. 8 (a) and (b). The tendencies of the responses

are the same as the tendencies of the Asymptotic Analysis, but there are some differences, because of the improved results in

Monte Carlo simulations due to the 1000 replications used.

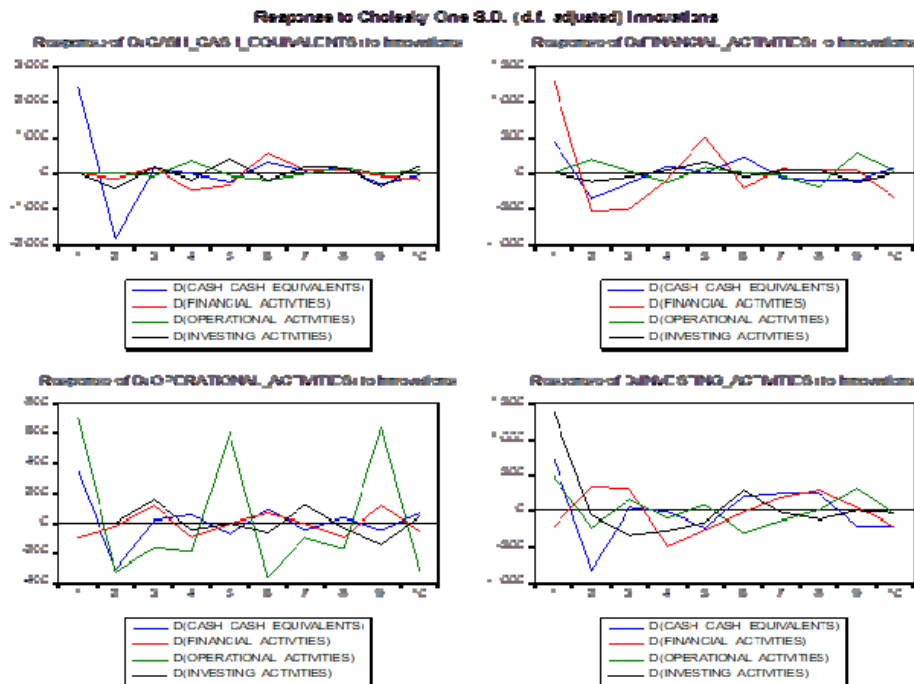


Fig. 10 The summary view of the treated cases for the Edgeworth Approximation model and Monte Carlo for the response of the series

In Fig. 10 is presented the summary view of the treated cases for the Edgeworth Approximation model and Monte Carlo for the response of the series after the increase of one Standard error of the other influencing series.

Model: Bootstrap in VAR model  
Sample (adjusted): 8 114  
Solve Options:  
Static-Stochastic Simulation  
Solver: Broyden  
Max iterations = 5000, Convergence = 1e-08  
Requested repetitions = 1000, Allow up to 5 percent failures  
Solution does not account for coefficient uncertainty in linked equations  
Track endogenous: mean, standard deviation, 95% confidence interval

Calculating Bootstrap Innovations  
Sample: @ALL  
107 residuals per equation calculated for bootstrap

Scenarios: Baseline & Baseline  
Solve begin 14:42:02  
Repetitions 1-200: successful 14:42:02  
Repetitions 201-400: successful 14:42:03  
Repetitions 401-600: successful 14:42:03  
Repetitions 601-800: successful 14:42:03  
Repetitions 801-1000: successful 14:42:03  
Solve complete 14:42:03  
1000 successful repetitions, 0 failure(s)

Fig. 11 Bootstrap in VAR model

In Fig. 11 are presented the results after the application of the Bootstrap Version that is related to the Baseline model that is used for the Stochastic relations, that aims the consideration of the errors done during the estimation of the VAR Model. For the Bootstrap Model are used 1000 replications, 95% confidence interval, 5% allowed failure.

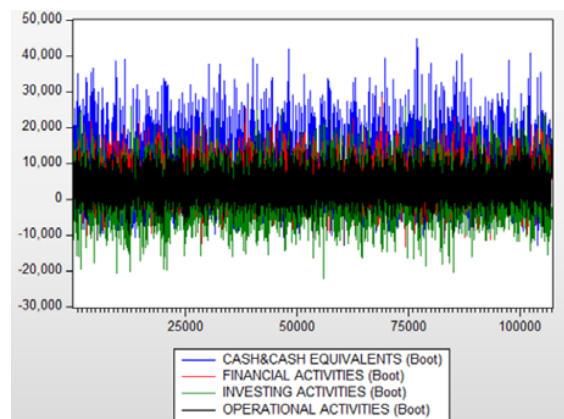


Fig. 12 The graphic where are overlapped the responses of all series

Fig. 12 presents the graphical view of the overlapped results of the responses for the four series of this study, after the particular shocks in the standard errors of the other influencing series. This visualized form could be even in the cases of Edgeworth Approximation and Monte Carlo Simulations that in fact, were presented in multiple graphics for the study's effect. So, we can conclude that The Edgeworth Approximation, Bootstrap Version and the Monte Carlo Simulations are the best methods that optimize the models that contain many variants. In concrete terms, our VAR model that is created based on the series created for the Cash & Cash Equivalents, Financial Activities, Operational Activities and

Investing Activities from the Financial Reported Information in the official site of Coca Cola, has optimized results based on the application of the three methods: Edgeworth, Bootstrap, Monte Carlo, that essentially have almost the same distribution.

## REFERENCES

- [1] A. Bose, Edgeworth Correction by Bootstrap in Autoregressions, The Annals of Statistics, Vol.16, No.4 (1988), 1709-1722.
- [2] A. Stuart and K. Ord, Distribution theory, vol.1. Oxford University press Inc., New York, USA, 1994, 74-162.
- [3] C. Sims, Macroeconomics and Reality, Econometrica, 48 (1980), 1-48.
- [4] D. Edgerton, and G. Shukur, Testing autocorrelation in a system perspective, Econometric Reviews, 18(1999), 343-386.
- [5] F. Götze and C. Hipp, Asymptotic expansions for sums of weakly dependent random vectors, Z. Wahrsch. Verw.Gebiete, 64 (1983), 211-239.
- [6] G.Casella and R. L. Berger, Statistical Inference, second edition. The Wadsworth Group, USA (2002), 83-87.
- [7] H. Lütkepohl, Asymptotic distributions of impulse response functions and forecast error variance decompositions of vector autoregressive models. The Review Of Economics And Statistics, 72(1990), 116-125.
- [8] H. Lütkepohl, Introduction to Multiple Time Series Analysis, New York: Springer-Verlag, 1991.
- [9] H. Lütkepohl, New Introduction to Multiple Time Series Analysis, New York: Springer-Verlag, 2007.
- [10] P. Hall, The Bootstrap and Edgeworth Expansion. Springer-Verlag, New York, USA, 1992.
- [11] R. Basna, Edgeworth Expansion and Saddlepoint Approximation for Descrete Data with Applications in Chance Games. Linnéuniversitetet, 2010.
- [12] R. N. Bhattacharya and J. K. Ghosh, on the Validity of the Formal Edgeworth Expansion, The Annals of Statistics, Vol.6, No.2, (1978), 434-451
- [13] R. W. Butler, Saddlepoint Approximation with Applications. Cambridge University Press, New York, USA (2007), 145.
- [14] S. Johansen, Likelihood-based Inference in Cointegrated Vector Autoregressive Models, Oxford: Oxford University Press, 1995.A. H. Welsh, Asymptotically Efficient Estimation of the Sparsity Function at a Point, Statistics & Probability Letters, 6 (1988), 427-432.
- [15] EViews Help: Multiple Equation Analysis
- [16] [http://fbemoodle.emu.edu.tr/pluginfile.php/47097/mod\\_resource/content/1/Workshop%204%20-%20Part%202%20-%20Advanced%20Time%20Series%20Econometrics%20with%20EViews.pdf](http://fbemoodle.emu.edu.tr/pluginfile.php/47097/mod_resource/content/1/Workshop%204%20-%20Part%202%20-%20Advanced%20Time%20Series%20Econometrics%20with%20EViews.pdf)