SOME CHARACTERISTICS OF THE FINANCIAL DATA SERIES

Ph.D Senior Lecturer Gheorghe SĂVOIU
University of Piteşti

Abstract

This paper attempts to delineate from a theoretical of view the financial data series relative to other statistical data, starting from the financial econometrics' models and from the resulting features of the specific descriptive statistics' analysis of these characteristic series. From the analysis of these financial data during either very short and short or medium periods of time or from the information provided by the website of the Bucharest Stock Exchange (BVB), the trend of great values of kurtosis or eccentricity and skewness or asymmetry of series appears as a characteristic tendency. During a long period of time, between 1920 and 2008, this tendency seems to be more relevant, being confirmed by an excerpt from the author's earlier paper written in 2009, concerning the statistical Dow Jones Industrial Average Index (DJIA Index). The skewness, kurtosis and normality of data distribution analysis, using Jarque Bera test, along with the identification of residual autocorrelation or serial correlation in the presence of significant residual values and heteroskedasticity are the major evaluated aspects. Finally, the author investigates the optimal way to ensure statistical comparability inflationary and deflationary method for financial series of data, and offers a solution to the selection of the appropriate indicator from the categories of the absolute values, absolute variation of the absolute values and the relative variation of the absolute values, expressed by percentages, with the finding of the latter alternative as the best alternative in the world of financial modelling of the economic and financial processes and phenomena.

Keywords: descriptive statistics, skewness, kurtosis, normal distribution, heteroskedasticity, residual autocorrelation, Eviews.

***

In practical economic-financial modelling it appears that the data series extracted from the financial markets reveal two important general characteristics distinguishing them firmly against all other data series. The first characteristic feature is the fact that those financial data series are more *arched* than usual economic series (they have a higher degree of *eccentricity* specific to financial phenomena dominated by risk and uncertainty, with is a larger deviation of extreme values from their average, along with a greater number of values cantered around the mean value, with the consequent increase in the number of statistical units placed closer to the central trend indicator (hence both a negative signal of the relative magnitude determined as to the extremes, and a direct and positive consequence linked to greater representation of the central tendency value, a better or firmer capacity of substitution the financial series by its average value).

In the entire family of indices of the Bucharest Stock Exchange (BSE or BVB) used below as an illustration of these features, we find that much more arching than in a usual series according to the values of the as vaulting or eccentricity coefficient, which are much higher than 3, even for a few days, when the exchange stock operated, between the 1st of January, 2013 and the 28th of February, 2013 (over a period of only 40 days):

Descriptive Statistics of BSE (BVB) Indices (1.01.2013 - 28.02.2013)

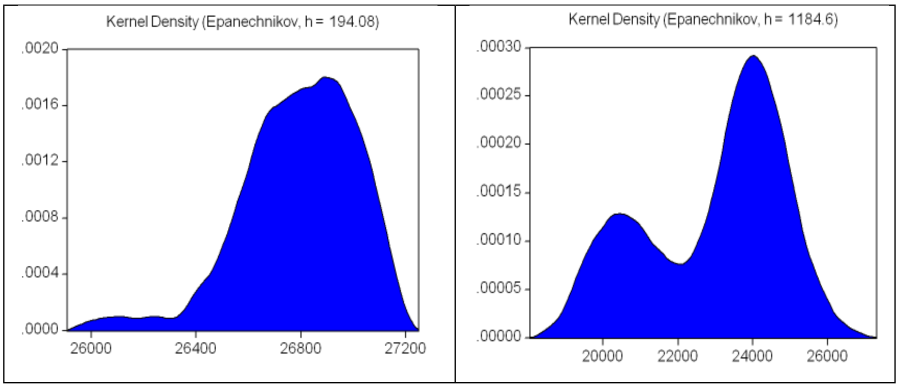
	BET	BET-C	BET-FI	BET-XT	BET-NG
Mean	5554.027	3022.619	26791.35	522.1565	676.1390
Median	5568.175	3037.720	26797.88	523.3400	678.7300
Maximum	5689.890	3085.720	27052.03	531.6500	708.4900
Minimum	5328.330	2877.000	26103.62	501.7300	645.3300
Std. Dev.	70.03657	45.40596	206.2881	5.840841	16.69446
Skewness	-0.814048	-1.443300	-1.015930	-1.227523	-0.070979
Kurtosis	4.312560	4.711406	4.391979	5.536763	2.020169

Source: http://www.bvb.ro/TradingAndStatistics/Evolution.aspx

Note: BET from BUCHAREST EXCHANGE TRADING

For a perfectly symmetrical or normal distribution, the value of asymmetry of the series of (financial) data or skewness is always “zero”. Overall, empirical data series are slightly or moderately asymmetric, as evidence that financial market on certain days recorded e.g. quotations that are larger or very large, or smaller and very small in relation to the previous day (defining positive and negative asymmetry). Another specific case on financial markets, in general, was represented by the negative value of skewness, indicating a negative or leftward asymmetry. Vaulting or eccentricity are defined as normal by a *kurtosis* equal to the value of 3, but specifically manifests in the (financial) data series when they show a degree of *vaulting* or excess distribution in the series of (financial) data, rather than a degree of *flattening*, according to growing values of *kurtosis* by temporal extension, and graphic distribution becomes higher (*leptokurtic*).

The sharper graphical vaulting and *kurtosis* increase from a 40 days (left) to 760 days (right) series of BSE (BVB) data



Source: http://www.bvb.ro/TradingAndStatistics/Evolution.aspx

Software used: Eviews

The second major aspect, which can even be considered the first in the light of the its econometric modelling impact, it is derived from the analysis of the normality of the distributions of the data series in the financial markets. These financial data series do not follow a normal distribution, as generally required and accepted in modelling, and according to the modelling assumptions issued, or as specified in the theoretical models. To illustrate the validity of this second feature, normality tests are achieved for the same data series of BSE indices. Testing the normality of distributions for financial statistical data highlights a majority trend of abnormality in keeping with the values resulting from the Jarque Bera test, which is increasing with the expansion of the data from only 40 days with quotations, to 250 days or 760 days with quotations.

The values resulting from the Jarque Bera test for the data of BSE (BVB) indices (1.01.2013-28.02.2013)

The Series of 40 days	BET	BET-C	BET-FI	BET-XT	BET-NG
Jarque-Bera	7.289178	18.76895	10.11010	20.77070	1.633702
Probability	0.026132	0.000084	0.006377	0.000031	0.441821

Source: http://www.bvb.ro/TradingAndStatistics/Evolution.aspx

- to 250 days with quotations in 2012:

The values resulting from the Jarque Bera test for the data of BSE (BVB) indices (1.01.2012-31.12. 2012)

The series of 250 days	BET	BET-C	BET-FI	BET-XT	BET-NG
Jarque-Bera	7.251069	32.18024	22.17329	10.34532	28.75489
Probability	0.026635	0.000000	0.000015	0.005669	0.000001

Source: http://www.bvb.ro/TradingAndStatistics/Evolution.aspx

or to 760 days with quotations (the 2010 -2012 financial data series), when none of the index series is otherwise than abnormally distributed, and the trend of abnormality of financial data series is general:

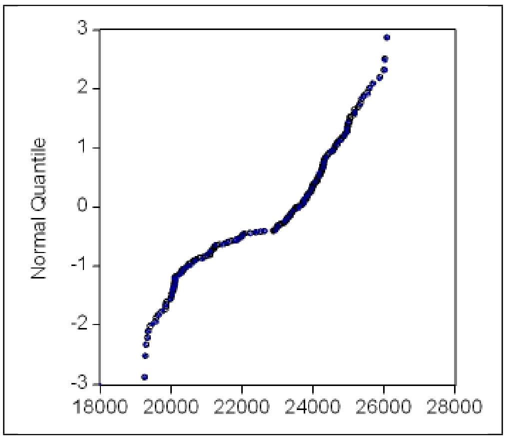
The values resulting from the Jarque Bera test for the data of BSE (BVB) indices (1.01.2010-31.12. 2012)

The series of 760 days	BET	BET-C	BET-FI	BET-XT	BET-NG
Jarque-Bera	22.89363	41.90127	53.06696	18.03000	53.30135
Probability	0.000011	0.000000	0.000000	0.000122	0.000000

Source: http://www.bvb.ro/TradingAndStatistics/Evolution.aspx

A simple method for comparing two distributions graphically, that is a normal theoretical series and another empirical one (the BET-FI series of 250 terms). In the Eviews package a Quanta Quanta-graphs (QQ-plots) is used. It can be noted that the empirical distribution BET-FI 250 is not normally distributed, the resulting graph 3 being clearly not focused on the first bisector, but completely different.

Quantile – quantile graphs for BET–FI 250



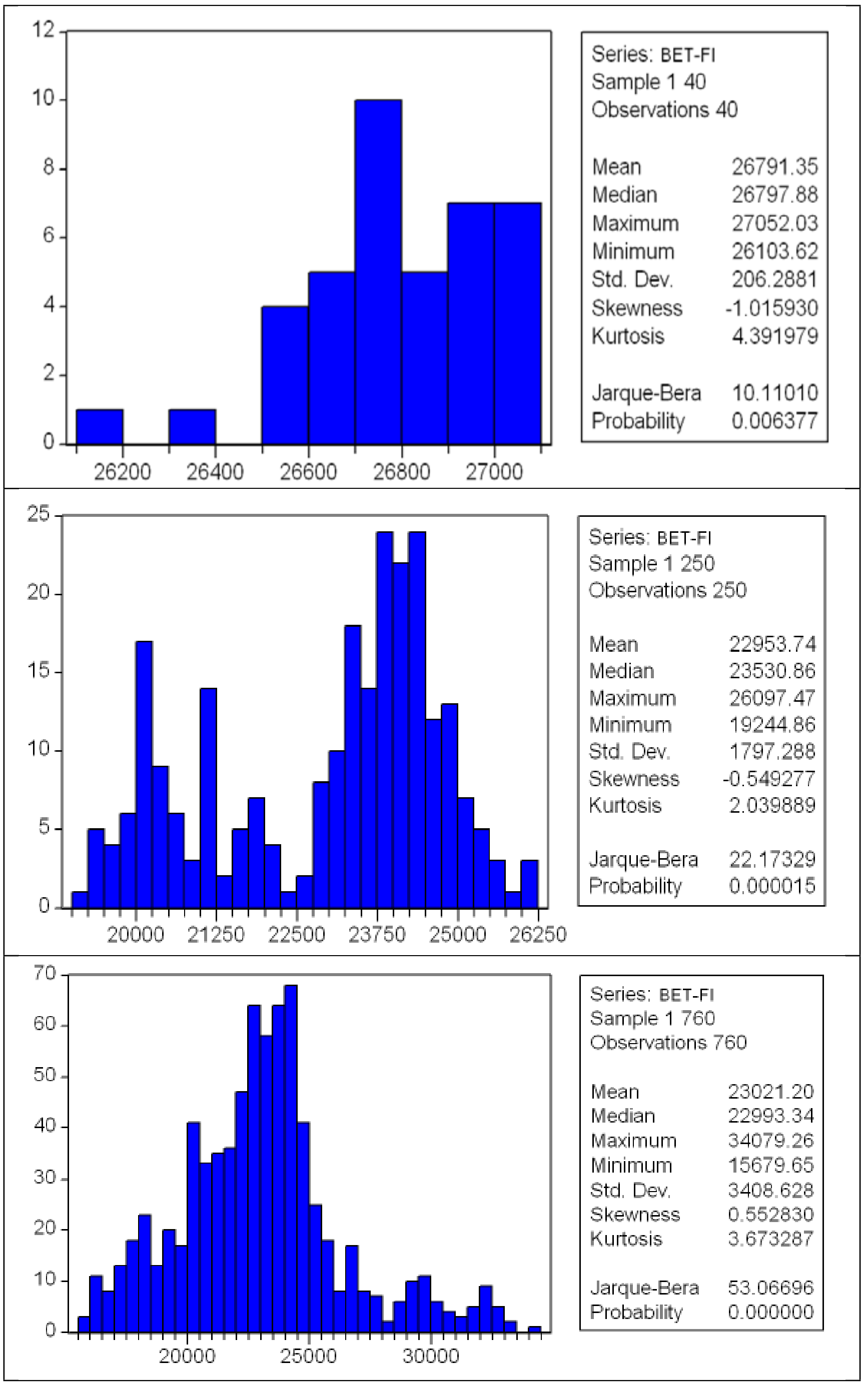
Source: http://www.bvb.ro/TradingAndStatistics/Evolution.aspx

Two other important issues for modelling financial data sets are generated by two delicate problems of model construction:

- the financial series show *serial correlation* of residues, or autocorrelation;

- the financial series also present obvious *heteroskedasticity* by their dispersion which is not uniform across all the range of temporal analysis and varies in amplitude (which is defined as *heteroskedasticity* – in English, *different variance* – from Greek *hetero* “different” and *skedasis* “scattering” or “dispersion”) and violates the fundamental assumption of homoskedasticity of the method of the smallest squares.

The histograms and successive descriptive statistics, in the 40-, 250- and 760 day series, with quotations of BET-FI, anticipating heteroskedasticity.



Series: BET-FI
Sample 1 40
Observations 40
Mean 26791.35
Median 26797.88
Maximum 27052.03
Minimum 26103.62
Std. Dev. 206.2881
Skewness -1.015930
Kurtosis 4.391979
Jarque-Bera 10.11010
Probability 0.006377
Series: BET-FI
Sample 1 250
Observations 250
Mean 22953.74
Median 23530.86
Maximum 26097.47
Minimum 19244.86
Std. Dev. 1797.288
Skewness -0.549277
Kurtosis 2.039889
Jarque-Bera 22.17329
Probability 0.000015
Series: BET-FI
Sample 1 760
Observations 760
Mean 23021.20
Median 22993.34
Maximum 34079.26
Minimum 15679.65
Std. Dev. 3408.628
Skewness 0.552830
Kurtosis 3.673287
Jarque-Bera 53.06696
Probability 0.000000

Time in the financial series generates greater uncertainty of econometric modelling, because it significantly increases or decreases the correlations of the data sets that characterize the financial markets, as reflected in the three correlation matrices below, which are used as an illustration:

Matrix of correlation of BSE (BVB) Indices (1.01.2013 - 28.02.2013)

40 days	BET	BET-C	BET-FI	BET-XT	BET-NG
BET	1.000000	0.715253	0.707588	0.982972	0.191508
BET-C	0.715253	1.000000	0.462402	0.755629	0.804445
BET-FI	0.707588	0.462402	1.000000	0.796851	0.037469
BET-XT	0.982972	0.755629	0.796851	1.000000	0.254172
BET-NG	0.191508	0.804445	0.037469	0.254172	1.000000

Matrix of correlation of BSE (BVB) Indices (1.01.2012-31.12.2012)

250 days	BET	BET-C	BET-FI	BET-XT	BET-NG
BET	1.000000	0.884537	0.857169	0.973494	0.760133
BET-C	0.884537	1.000000	0.663592	0.863326	0.956765
BET-FI	0.857169	0.663592	1.000000	0.938399	0.605099
BET-XT	0.973494	0.863326	0.938399	1.000000	0.782096
BET-NG	0.760133	0.956765	0.605099	0.782096	1.000000

Matrix of correlation of BSE (BVB) Indices (1.01.2010-31.12.2012)

760 days	BET	BET-C	BET-FI	BET-XT	BET-NG
BET	1.000000	0.950159	0.719589	0.948650	0.878631
BET-C	0.950159	1.000000	0.585660	0.886779	0.956963
BET-FI	0.719589	0.585660	1.000000	0.884891	0.443029
BET-XT	0.948650	0.886779	0.884891	1.000000	0.772484
BET-NG	0.878631	0.956963	0.443029	0.772484	1.000000

The analysis of the data series for very long periods has led to the efficient market hypothesis (EMH), which requires a clarification of the market efficiency hypothesis. The efficient market is actually a paradox. If every investor believes that the market is efficient, and then the market will no longer be so, no one will buy shares, because nobody can earn more. The efficiency of a market also depends on how many investors believe it is effective. *As a logical consequence, markets are not, and cannot be completely efficient, nor can they be entirely or completely ineffective*. Further evidence of the specificity of financial series, offered by the informational efficiency of stock markets and capital markets, gradually turned more rational and more symmetrical, is provided by the descriptive analysis of one of the best-known indexes, the Dow Jones Industrial Average (DJIA) Index, one of the oldest, extended, rational and symmetric financial markets in the world.

The features of the financial series defined are maintained for the *model of random distribution of DJIA values* over eight decades, namely the excess of vaulting (very high values of *kurtosis*) and the negative dominant of the value of *skewness.*

The descriptive statistics of the Dow Jones Industrial Average (DJIA) Index

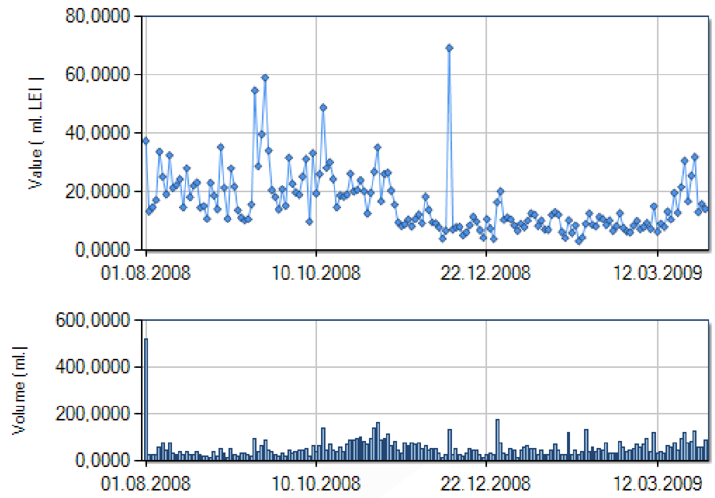
Interval	Mean	Std. Dev.	Kurtosis	Skewness	H-lbqtest	H-chi2
1930-1939	-0.0196	2.0232	8.0295	0.298	1	1
1940-1949	0.0105	0.82	12.1952	-1.1305	1	1
1950-1959	0.0487	0.671	10.1739	-0.9169	1	1
1960-1969	0.0067	0.6533	8.5218	-0.0481	1	1
1970-1979	0.0006	0.9263	4.7826	0.2742	1	1
1980-1989	0.0487	1.1561	103.5361	-4.3501	1	0
1990-1999	0.0553	0.892	8.1889	-0.4086	1	1
2000-2008	-0.0087	1.2907	11.5553	0.0038	1	1
1928-2008	0.0175	1.1602	28.44	-0.6147	1	0

Source: Săvoiu, G., Andronache, C., *Rationality and information symmetry EMH and DJIA* [Raţionalitate şi simetrie informaţională EMH şi DJIA], Romanian Statistical Review, supplement no. 1 2009. p. 221.

Overall empirical data series are slightly or moderately asymmetric as evidence or as a proof that on financial market during certain periods exchange prices were higher and even very high or lower and very low in relation to the previous day (defining moderate or pronounced asymmetry, either positive or negative). Another problem arises here, that is the correct definition of a financial variable BET as: a) the absolute value of the index; b) the absolute variation of the absolute value of the index; c) the percentage change of the absolute value of the index.

This aspect multiplies the variables that will be investigated by a statistician and sometimes creates complications in the general approach of the statistical thinking if the information must be reunited and tested with Jarque – Bera test.

The evolution of the Bucharest Stock Exchange (BSE or BVB)
1.08.2008-1.04.2009



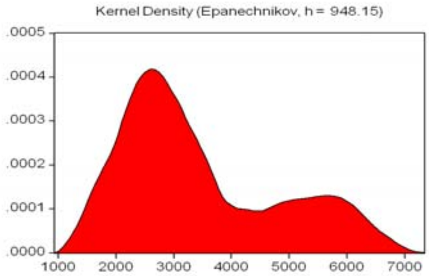
Source: http://www.bvb.ro/TradingAndStatistics/Evolution.aspx?s=bse&m=lei&c=dd&d1=3/14/2013&d2

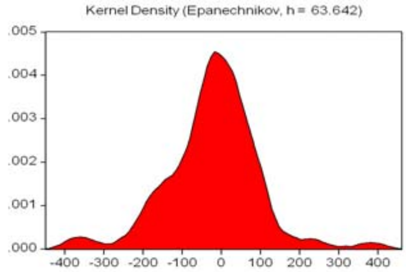
From the descriptive statistics of BET, BET (var.), and BET (var. %):

The descriptive statistics of BET, BET (var.) şi BET (var. %)
1.08.2008-1.04.2009

Data Sample: 1 162	BET	BET (Var.)	BET (Var. %)
Mean	3432.968	-21.73673	-0.494321
Median	2941.465	-14.91500	-0.455000
Maximum	6380.760	395.2800	10.62000
Minimum	1887.140	-384.2400	-12.29000
Std. Dev.	1320.394	115.3098	3.377231
Skewness	0.833171	-0.070669	-0.092864
Kurtosis	2.386173	5.164521	4.671634
Jarque-Bera	21.28599	31.75960	19.09476
Probability	0.000024	0.000000	0.000071
Sum	556140.9	-3521.350	-80.08000
Sum Sq. Dev.	2.81E+08	2140713	1836.316
Observations	162	162	162

as well as from the distribution's graphs of these variables:



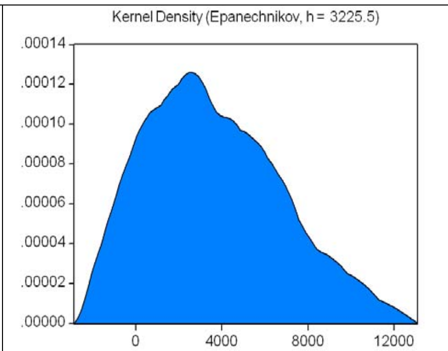


Software used: Eviews

it can be seen that annual BET (var.%) is the most adequate indicator for analysis on long term, and thus the relative and annual values (Graph no. 11) are normally distributed and the most used indicators in financial series.

BET annual values between 1997 and 2012

Year	BET
1997	757.86
1998	377.63
1999	448.53
2000	544.72
2001	754.85
2002	1659.06
2003	2171.88
2004	4364.71
2005	6586.13
2006	8050.18
2007	9825.38
2008	2901.10
2009	4690.57
2010	5268.61
2011	4336.95
2012	5149.56



BET Descriptive Statistics

Sample: 1997 2012 BET	
Mean	3617.982
Median	3619.025
Maximum	9825.380
Minimum	377.6300
Std. Dev.	2910.806
Skewness	0.585844
Kurtosis	2.394089
Jarque-Bera	1.159988
Probability	0.559902
Sum	57887.72
Sum Sq. Dev.	1.27E+08
Observations	16

Software used: Eviews

Providing statistical comparability of financial data series is another aspect having serious implications in economic and financial modelling, unless properly addressed. Investigation of the economic-financial phenomenon involves the same temporal, spatial and structural universe of the phenomena modelled, which involves using both econometric models constructed from indicators of value, and models exclusively made of absolute and relative indicators (especially indices and rates), as well as modelling focused on indicators rendering the implications of changes in the structure of the financial phenomenon. Financial modelling also requires a certain initial option as to the length of the data series (according to criteria of availability, comparability and relevance), for proper periodization (here in the sense of cutting out a pretty recent time interval as current impact, derived from prediction or forecasting requirements of modelling, for ensuring a comparable level of volume, impact and intensity of the flows analyzed). Some phenomena of European regional impact may require abandoning expression of value in lei for Euros, which is sufficiently justified in terms of the dominant volume of the financial phenomenon analyzed in the euro, though the need to compare the same global phenomenon can send to the solution of quantifying by means of U.S. dollars.

The option for an appropriate data source or choosing between the databases of the World Bank, may occur to the detriment of the data from Eurostat or the INS, which provide partial coverage (i.e. national and European) of the economic and financial phenomenon. The issue of the comparability of the financial phenomenon over longer periods may also require priority selection of a base year, giving up the years where the methodologies are not fully comparable (e.g. 1990 cannot be integrated in the study of the investment phenomenon in Romania, being a year of fixing the value of aggregate national data). Statistical comparability also demands ensuring a coherent process of transforming all values in current prices by assessments in comparable prices (in the phenomena of updating financial values by means of inflation, it is preferred to use a central year of the data series).

Conclusions

A descriptive statistical investigation of financial data series emphatically points out that they are more arched / vaulted than the other, regular economic series (they have a higher degree of *eccentricity* specific to the financial phenomena dominated by risk and uncertainty) often possess negative values of skewness, do not have a normal distribution, as generally accepted in the modelling requirements and according to modelling assumptions emitted, or as specified in the theoretical model, show *serial correlation* of residue or autocorrelation, and *heteroskedasticity*, and ensuring statistical comparability of financial data series is another aspect with serious implications in economic-financial modelling, if not properly addressed and solved.

Time in the financial series generates greater uncertainty of econometric modelling, because it significantly increases or decreases the correlations of the data sets that characterize the financial markets.

References

- Andrei, T., Bourbonnais, R., *Econometrie*, Ed. Economică, Bucureşti, 2008.
- Andrei, T., Stancu, S., Iacob, A.I., Tuşa, E., *Introducere în econometrie utilizând EViews*, Ed. Economică, Bucureşti, 2008.
- Clocotici,V., *Introducere în statistica multivariată,*Ed.Universităţii, A.I. Cuza, Iaşi, 2007.
- Codirlaşu, A. *Econometrie aplicată utilizând EViews 5.1 2007.* http://www.dofin.ase.ro/ acodirlasu/lect/econmsbank/econometriemsbank2007.pdf.
- Greene, W. H., *Econometric Analysis*, 4th ed., Prentice Hall International, 2000.
- Gujarati, D.N.,Porter, D.C., *Basic econometrics*, Boston: McGraw Hill, 5th Edition, 2009.
- Jaba E. (coord.), *Econometrie aplicată*, Ed. Univ. Al. I. Cuza, Iaşi, 2008.
- Onicescu, O., *Econometrie şi statistică informaţională*, Comunicări DCS, Bucureşti,1983.
- Pecican, E., *Econometrie ... pentru economişti*, Ed. Economică, Bucureşti, 2004.
- Săvoiu, G., *Econometrie* ,Ed. Universitară, Bucureşti, 2011
- Săvoiu, G., (Modelarea economico – financiară : Gândirea econometrică aplicată în domeniul financiară, Ed. Universitară, Bucureşti, 2013
- Vogelvang, B., *Assignments for:"Econometrics, Theory and Applications with EViews"* Vrije Universiteit Amsterdam, Department of Econometrics, 2005,Version 01.
- Voineagu,V., Ţiţan,E., Şerban,R., Ghiţă,S., Tudose,D., Boboc,C., Pele,D., *Teorie şi practică econometrică*, Ed Meteor Press, Bucureşti, 2007.
- ***http://www.eviews.com
- ***http://www.insse.ro
- ***http://www.bvb.ro/TradingAndStatistics/