Liet. mat. rink. LMD darbai, **50**, 2009, 190–195

Factor analysis of the Lithuanian equity market indices

Svetlana DANILENKO
Vilnius Gediminas Technical University
Saulėtekio 11, 10223 Vilnius, Lithuania
e-mail: svetlana.danilenko@fm.vgtu.lt

Abstract. Statistical measures that can reproduce the state of the stock market and the tendencies of its change dynamics are the stock indexes. Having in mind the more complicated state of the finance system it is important to answer the question of what impacts the fluctuations of the stock prices. The article discusses various factors that impact the fluctuations of the Lithuanian stock index $OMXV$; also stock index factor analysis is performed. Factors are determined using the main components method.

Keywords: stock market, factor analysis.

Introduction

Instability of the share prices is influenced by many factors: both clearly measurable quantitative (microeconomic, macroeconomic and other indicators) and clearly not defined qualitative aspects (social, political, psychological and other). Arbitrage theory describes that a limited risk component quantity exists [6]. Ross suggested that the components represent the fundamental macroeconomic and financial indicators. Which means that the existing indicators fairly accurately describe the economic situation, which had formed in the country at a certain point in time, and can be analysed as factors influencing the share prices. It has been shown that at least three factors influence the asset prices [5].

The article analyses the influence of sector indices on the Lithuanian share index $OMXV$ return during years 2000–2008, and factor analysis of the share indices is performed. The factor analysis method is one of the most commonly known classical multi-measure methods [2,7,8]. The purpose of factor analysis – to replace the set of characteristics of the observed phenomenon by a combination of several factors while losing minimal information.

Research method

Variable description. The article analyses the impact of sector indices on the Lithuanian share index $OMXV$ return during the years 2000–2008. With the help of factor analysis method factors are determined (the main components method is used) that describe the main Lithuanian share sectors, using the Granger causality method the causation relationship is determined.

Sector index indicators researched: energy ($Energ$), commodity ($Comm$), production ($Prod$), consumption of goods and services ($Consum$), consumption of essential goods and services ($Esset$), healthcare ($Health$), finance ($Financ$), information technology (IT), telecommunication services ($Telecom$), utility services ($Utilit$).

Normally share price changes are researched since they reflect the changing economic environment. Let us assume $P(t)$ share index at time. The return for the same period is the share index change logarithm $r(t)$:

$$r(t) = \ln\left(\frac{P(t)}{P(t-1)}\right). \tag{1}$$

The Lithuanian share index return is calculated according to the indicated formula $rOMXV$.

Preparation of data for the factor analysis to be performed. The preparation of data for the factor analysis had been performed in two stages [4].

First stage – ensuring stationarity. The process is stationary if its mean and dispersion are constant values, i.e., they do not depend on the time variable. The existing data is not stationary, it needs to be transformed. One of the methods applied is to differentiate the process, i.e., every value of the time series is substituted with the difference of the existing and preceding values. To achieve stationarity second-order differentiation is applied to the transformation of the primary data.

Second stage – standardisation of the time series. Since the existing data is not standardised it needs to be transformed. The data standardisation is performed by deducting the mean of the time series and dividing by the variance, thus the mean of all of the standardised data is zero, and the variance – one. The data must be standardised so that the variables with a large deviation would not dominate and would not alter the results.

Description of the mathematical model. Let us assume that we are observing s variables $X_1, X_2, \ldots, X_s$. The model is based on the assumption that behaviour of every variable X_i is caused by m general latent factors $F_1, F_2, \ldots, F_m$ and a specific latent factor e_i. There are less latent factors than there are variables, i.e., $m < s$. Let the variable X_i be linearly dependent on the factors. The mathematical model is:

$$X_i = \sum_{j=1}^{m} \lambda_{ij} F_j + e_i. \tag{2}$$

Multiples λ_{ij} are called the factor weights.

Externally the factor analysis model resembles linear regression – knowing F_j and λ_{ij} values we could predict the values of X_i. However, the purpose of factor analysis is the opposite – only X_i values are known, and we want to estimate the general factors F_j.

When applying factor analysis the similarity in observed variables is sought. When the variables are not correlated the factor analysis would be purposeless, therefore, it needs to be ensured that the variables observed are correlated. From existing data only the correlated should be selected, and the uncorrelated should be eliminated from the factor analysis data list.

The primary factor analysis data – observation correlations (or covariance) matrix. It can be seen from its form which variables are correlated.

Suitability of data for factor analysis can be evaluated using the Kaizer–Meyer–Olkin (KMO) adequacy coefficient. It is the empirical correlation coefficient and partial correlation coefficient magnitude comparative index

$$KMO = \frac{\sum\sum_{i \neq j} r_{ij}}{\sum\sum_{i \neq j} r_{ij} + \sum\sum_{i \neq j} \overline{r}_{ij}}, \tag{3}$$

where r_{ij} is correlation coefficient of X_i and X_j, $\overline{r}_{ij}$ – X_i and X_j partial correlation coefficient. If KMO value is low then the factor analysis of the observed variables is inconclusive. When KMO value is less than 0.6, then the correlation of variable pairs cannot be explained by other variables, and the factor analysis of the observed variables is not acceptable.

Suitability measure of the observations of every variable can be calculated using the following formula

$$MSA_i = \frac{\sum_{i \neq j} r_{ij}}{\sum_{i \neq j} r_{ij} + \sum_{i \neq j} \overline{r}_{ij}}. \tag{4}$$

The variable which has the smallest MSA value is removed from the primary variables list and then the measure of suitability of data for factor analysis is calculated KMO. The procedure is repeated until KMO value exceeds 0.6.

Granger causality test. When using the Granger causality test [1] we will research the observed variables' causation relationship. Granger causality test for time series is based on the assumption that: if x causes y then before y changes, the x must change, but not vice versa.

In other words, two conditions must be met:

(a) x must input a statistically significant contribution to the prediction of y;

(b) y must not input a statistically significant contribution to the prediction of x.

When verifying the Granger causality the following regression equations are formed:

$$y_t = \alpha_0 + \sum_{i=1}^{m} \alpha_i y_{t-i} + \sum_{i=1}^{m} \beta_i x_{t-i} + \varepsilon_t, \tag{5}$$

$$x_t = \alpha_0 + \sum_{i=1}^{m} \alpha_i x_{t-i} + \sum_{i=1}^{m} \beta_i y_{t-i} + u_t, \tag{6}$$

where ε_t and u_t are uncorrelated random errors.

The null hypothesis that the coefficients are statistically significant is tested for every equation

$$\beta_1 = \ldots = \beta_m = 0. \tag{7}$$

Let us highlight that the fact that x causes y shows only that earlier values of x explain further values of y, i.e., shows the possibility of causation. If the null hypothesis that x does not cause y is not rejected, it means x does not cause y [3].

We select 5% confidence intervals for testing the null hypothesis.

Review of results

All the variables are used for composing the initial factor model. The suitability of data is verified using KMO. Since KMO is 0,692 all the variables are left in the model.

Tables 1–3 contain the summarising results that are based on the main components method. From Table 1 it can be seen that three first components explain 75.91 % of the total variance:

- variance of the first component is 4.72, it explains 47.20 % of the researched sector index general variance,
- variance of the second component is 1.59, it explains 15.86 % of the researched sector index general variance,
- variance of the third component is 1.28, it explains 12.84 % of the researched sector index general variance.

The general position of the Lithuanian sector index market can be represented using three components. The first component describes the energy, production, consumption of goods and services, healthcare as well as finance sector indices. The second component describes the consumption of essential goods and services, telecommunication services and utility service sector indices. The third component describes the commodity and information technology sector indices.

The second stage analyses the causation relationship of the primary variables and the factor analysis method factors (F_1, F_2, F_3).

Table 1. The explained part of factor variation

	Main components									
	1	2	3	4	5	6	7	8	9	10
Value	4.72	1.59	1.28	0.86	0.53	0.36	0.27	0.21	0.14	0.04
Percentage	47.20	15.86	12.84	8.61	5.32	3.59	2.69	2.05	1.42	0.41
Cumulative percentage	47.20	63.07	75.91	84.51	89.84	93.43	96.12	98.17	99.59	100

Table 2. Transposed component matrix

	Main components		
	1	2	3
Energ	**0.832**	0.204	0.108
Comm	0.404	0.357	**0.701**
Prod	**0.745**	0.483	−0.157
Consum	**0.750**	0.466	−0.160
Essent	0.155	**0.844**	0.105
Health	**0.837**	−0.319	0.181
Financ	**0.804**	0.421	−0.018
IT	0.122	0.085	**−0.682**
Telecom	0.217	**0.756**	−0.275
Utilit	0.158	**0.809**	0.419

Table 3. Selection of factors

Main components		
1	2	3
$Energ$	$Essent$	$Comm$
$Prod$	$Telecom$	IT
$Consum$	$Utilit$	
$Health$		
$Financ$		

Table 4.Granger causality test results

Hypothesis	F-statistics	Level of importance (p value)
$Energ$ does not cause $rOMXV$	1.270	0.313
$Comm$ does not cause $rOMXV$	2.428	0.080
$Prod$ does not cause $rOMXV$	1.704	0.187
$Consum$ does not cause $rOMXV$	3.263	0.031
$Essent$ does not cause $rOMXV$	2.456	0.077
$Health$ does not cause $rOMXV$	3.959	0.015
$Financ$ does not cause $rOMXV$	3.935	0.015
IT does not cause $rOMXV$	0.281	0.887
$Telecom$ does not cause $rOMXV$	0.025	0.999
$Utilit$ does not cause $rOMXV$	2.161	0.109
F_1 does not cause $rOMXV$	3.469	0.025
F_2 does not cause $rOMXV$	0.731	0.581
F_3 does not cause $rOMXV$	0.920	0.471

After applying the Granger causality test it has been determined that the returns of the Lithuanian share indices $rOMXV$ are caused by: the consumption of goods and services, healthcare, finance sector as well as the first component expressed using the main component method (see Table 4).

Conclusion

For years the share return and the country fundamental indicator relationship has been researched in various academic articles and works. Different publications describe the share return relationship with the economic, financial and political indicators.

To research the impact of sectors to Lithuanian share market 10 main sector indices have been analysed.

It has been determined that the observed Lithuanian share market indices can be expressed using three factors.

With the help of Granger causality test the causation relationship has been discovered between the Lithuanian share index return and the sector share index indicators. It has been determined that Lithuanian share index return cause can be put down to three sector indices from the 10 used in this paper as well as one of the found factors.

References

1. C.W. Granger. Investigating casual relations by econometric methods and cross-spectral methods. *Econometrica*, 37: 424–438, 1969.
2. J.F. Hair, R.E. Anderson, R.L. Tatham, W.C. Black. *Multivariate Data Analysis*. Prentice-Hall, Upper Saddle River, NJ, 1998.
3. P. Jorion. *The Linkages Between National Stock Markets. The Handbook of International Financial Management.* Dow-Jones Irwin, Illinois, 1989.
4. M. Marcellino, J.H. Stock, M.W. Watson. Macroeconomic forecasting in the Euro area: country specific versus euro wide information. *European Economic Review*, 47:1–18, 2003.
5. R. Roll, S.A. Ross. An empirical investigation of the arbitrage pricing theory. *Journal of Finance*, 35:1073–1103, 1980.
6. S.A. Ross. The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13:341–360, 1976.
7. B.G. Tabachnick, L.S. Fidell. *Using Multivariate Statistics*. Allyn and Bacon, Boston, MA, 2001.
8. R.S. Tsay. *Analysis of Financial Time Series*: *Financial Econometrics*. Wiley, New York, NY, 2002.

REZIUMĖ

S. Danilenko. Lietuvos vertybinių popierių rinkos faktorinė analizė

Statistiniai rodikliai, įgalinantys perteikti akcijų rinkos būklę ir jos dinamikos kitimo tendencijas, yra akcijų indeksai. Atsižvelgiant į besikomplikuojančią pasaulio finansų sistemos būklę svarbu atsakyti į klausimą, kokie veiksniai daro įtaką akcijų kainų svyravimams. Straipsnyje nagrinėjami faktoriai lemiantys Lietuvos akcijų indekso $OMXV$ svyravimus, atliekama akcijų indeksų faktorinė analizė. Faktoriai nustatomi taikant pagrindinių komponenčių metodą.

Raktiniai žodžiai: vertybinių popierių rinka, faktorinė analizė.