Journal of Engineering and Applied Sciences 12 (19): 4922-4926, 2017
ISSN: 1816-949X

Peculiarities of Modelling of the Enterprise Investment Attractiveness in the Conditions of Multicollinearity of Predictors

[1]N.M. Yakupov, [2]G.M. Kvon, [1]M.F. Salahieva, [1]E.I. Kadochnikova and [1]A.V. Lelyukh
[1]Institute of Management, Economics and Finance, Kazan Federal University, 420008 Kazan, Russia
[2]Department of Economic Theory, Kazan National Research Technical University Named after A.N. Tupolev-KAI (KNRTU-KAI), 420111 Kazan, Russia

Abstract: The study presents an approach to assess the enterprise investment attractiveness based on the econometric modeling of return on assets. The researchers underline the key role of financial indicators in assessing investment attractiveness and propose a system of financial ratios of the enterprise-predictors of return on assets. In the conditions of collinearity of prognostic factors, the researchers offer to implement ridge regression which enables to obtain better prognostic characteristics to preserve reliability and informational value of the modelling. The researchers suggest tools to analyze how predictors of return on assets contribute to the assessment of investment attractiveness whose quality was tested using standard fisher and student tests and the standard error. The results of the empirical evaluations carried out using the Gretl Software confirmed their feasibility for potential investors, shareholders and owners in managing the use of capital effectively.

Key words: Investment attractiveness, principal components regression, return on assets, evaluation, Kazan, Russia

INTRODUCTION

Fearing global risks, investors now require more convincing evidence of the investment appeal of the company. Therefore, development of investment strategy requires a systematic approach to investment market research. In this regard, the relevance of research to develop scientifically sound and practically applicable methods of analyzing investment appeal of the industrial enterprises is increasing. Existing approaches to identify and differentiate factors and indicators of company's investment appeal as a whole can be divided into two main areas: in one case, the researchers identify only internal factors and indicators of enterprise's investment appeal, in other both internal and external. However, for potential investors the most manageable and informative factors are internal ones. According to Cokins (2009), Walsh (2011) and Altman (1968), the main internal factor of company's investment appeal is its financial standing which is a general characteristic of company's performance. The traditional tool to assess the investment appeal of the enterprise based on this factor is the analysis of its proportions (ratios) using its financial statements. The researchers (Khasanova *et al*., 2015; Kapler, 2000; Berestetsky and Sapotnitsky, 2004; Khavin, 2004) pointed out that the financial standing and accordingly, the investment appeal of the company is largely determined by the net return on assets in its possession. Return on assets (on total capital invested) shows whether the enterprise has the base to ensure high return on investor's equity. The company that does not have high values of net return on assets, is virtually unable to provide a sufficiently high level of return on investor's equity. When developing methodical approaches to evaluation of investment appeal, it seems appropriate to use multiple linear regression of the net return on assets in order to identify the most significant regressor. Econometric multivariate model of net return on assets will allow to quantify the impact of the individual regressors to identify the most significant of them, determine the "bottlenecks" in increasing the investment appeal and determining its indicators in the long term. The working hypothesis of the study is that the increase of the net return on assets is promoted by the growth of return on sales, turnover of current assets, current liquidity, increase in the ratio of payables to loan capital,

Corresponding Author: N.M. Yakupov, Institute of Management, Economics and Finance, Kazan Federal University, 420008 Kazan, Russia

the ratio of receivables and payables; while the decrease in the ratio of current liabilities to receivables, decrease in the ratio of debt capital and assets leads to an increase in the net return on assets.

MATERIALS AND METHODS

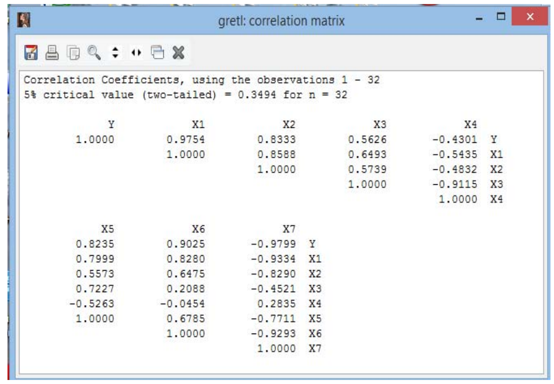
Using classical least squares, we have built the original model of net return on the enterprise's assets, calculated the inflation regression test VIF for each regressor, then we applied a ridge regression to estimate the parameters of multivariate linear regression model. Quality comparison of estimates of the models received was performed by standard Fisher test and t-test (Wooldridge, 2009; Hill *et al*., 2012; Samerkhanova and Kadochnikova, 2015) according to the standard error models and information criteria of Akaike and Schwarz. To argue the solution of the study hypothesis we used quarterly financial ratios for 2009-2016 of a leading chemical enterprise; they were obtained by calculations of data from "Spark" online system: Y, net return on assets, %; X1, net return on sales,%; X2, turnover of current assets; X3, current ratio; X4, the ratio of current liabilities to receivables; X5, the ratio of receivables and payables; X6, the ratio of payables to loan capital,%; X7, the ratio of debt capital and assets. Modeling was developed using Gretl Software package (Adkins, 2014).

RESULTS AND DISCUSSION

Linear coefficients of pair correlation received by Gretl Software between the dependent variable and regressors as well as linear coefficients of correlation between factors is presented in Fig. 1.

The analysis of linear coefficients of pair correlation of regressors with the dependent variable-net return on assets (Ryxj) (Fig. 1), showed that the net return on assets has a close direct correlation with the net return on sales (Ryx1 = 0.975), turnover of current assets (Ryx2 = 0.833), the ratio of receivables and payables (Ryx5 = 0.824), the ratio of accounts payable and debt capital (Ryx6 = 0.903), it has a close inverse correlation with the ratio of debt capital and assets (Ryx7 = -0.980), regressors X3, current ratio, X4, the ratio of current liabilities to accounts receivable have a moderate correlation with the net return on assets.

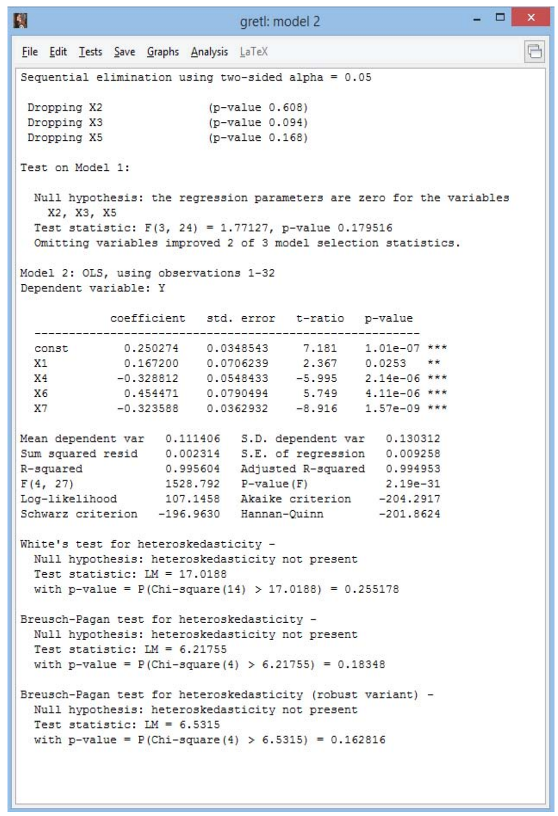
However, the linear coefficients of factor correlation (Rxixj) show the close correlation (collinearity) between regressors: Rh1x2 = 0.859; Rh3x4 = -0.912; Rh1x5 = 0.800; Rh3x5 = 0.723; Rx1h6 = 0.828; Rx1h7 = -0.933; Rx2h7 = -0.829; Rh6x7 = -0.929.



gretl: correlation matrix

```
Correlation Coefficients, using the observations 1 - 32
5% critical value (two-tailed) = 0.3494 for n = 32

          Y          X1          X2          X3          X4
     1.0000      0.9754      0.8333      0.5626     -0.4301  Y
                 1.0000      0.8588      0.6493     -0.5435  X1
                             1.0000      0.5739     -0.4832  X2
                                         1.0000     -0.9115  X3
                                                     1.0000  X4

         X5          X6          X7
     0.8235      0.9025     -0.9799  Y
     0.7999      0.8280     -0.9334  X1
     0.5573      0.6475     -0.8290  X2
     0.7227      0.2088     -0.4521  X3
    -0.5263     -0.0454      0.2835  X4
     1.0000      0.6785     -0.7711  X5
                 1.0000     -0.9293  X6
                             1.0000  X7
```

Fig. 1: Matrix of linear coefficients of pair correlation



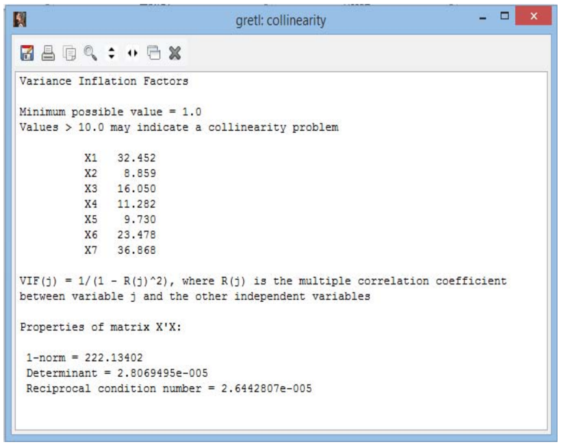
gretl: model 2

File Edit Tests Save Graphs Analysis LaTeX

```
Sequential elimination using two-sided alpha = 0.05

 Dropping X2               (p-value 0.608)
 Dropping X3               (p-value 0.094)
 Dropping X5               (p-value 0.168)

Test on Model 1:

  Null hypothesis: the regression parameters are zero for the variables
    X2, X3, X5
  Test statistic: F(3, 24) = 1.77127, p-value 0.179516
  Omitting variables improved 2 of 3 model selection statistics.

Model 2: OLS, using observations 1-32
Dependent variable: Y

             coefficient   std. error   t-ratio   p-value
  ---------------------------------------------------------
  const        0.250274    0.0348543     7.181    1.01e-07 ***
  X1           0.167200    0.0706239     2.367    0.0253   **
  X4          -0.328812    0.0548433    -5.995    2.14e-06 ***
  X6           0.454471    0.0790494     5.749    4.11e-06 ***
  X7          -0.323588    0.0362932    -8.916    1.57e-09 ***

Mean dependent var   0.111406   S.D. dependent var   0.130312
Sum squared resid    0.002314   S.E. of regression   0.009258
R-squared            0.995604   Adjusted R-squared   0.994953
F(4, 27)             1528.792   P-value(F)           2.19e-31
Log-likelihood       107.1458   Akaike criterion    -204.2917
Schwarz criterion   -196.9630   Hannan-Quinn        -201.8624

White's test for heteroskedasticity -
  Null hypothesis: heteroskedasticity not present
  Test statistic: LM = 17.0188
  with p-value = P(Chi-square(14) > 17.0188) = 0.255178

Breusch-Pagan test for heteroskedasticity -
  Null hypothesis: heteroskedasticity not present
  Test statistic: LM = 6.21755
  with p-value = P(Chi-square(4) > 6.21755) = 0.18348

Breusch-Pagan test for heteroskedasticity (robust variant) -
  Null hypothesis: heteroskedasticity not present
  Test statistic: LM = 6.5315
  with p-value = P(Chi-square(4) > 6.5315) = 0.162816
```

Fig. 2: OLS estimate of multiple regression of net return on assets for the full set of factors

In the presence of collinearity, for the selection of regressors into multiple regression, you must first assess multiple regression with the classical Least Squares (OLS) (Fig. 2) for full set of regressors and determine the regression inflation criterion (VIFj) to detect excess collinear regressors. Based on the received results shown in Fig. 2, we can write the original multiple linear regression model of the net return on assets:



gretl: collinearity

```
Variance Inflation Factors

Minimum possible value = 1.0
Values > 10.0 may indicate a collinearity problem

        X1   32.452
        X2    8.859
        X3   16.050
        X4   11.282
        X5    9.730
        X6   23.478
        X7   36.868

VIF(j) = 1/(1 - R(j)^2), where R(j) is the multiple correlation coefficient
between variable j and the other independent variables

Properties of matrix X'X:

 1-norm = 222.13402
 Determinant = 2.8069495e-005
 Reciprocal condition number = 2.6442807e-005
```

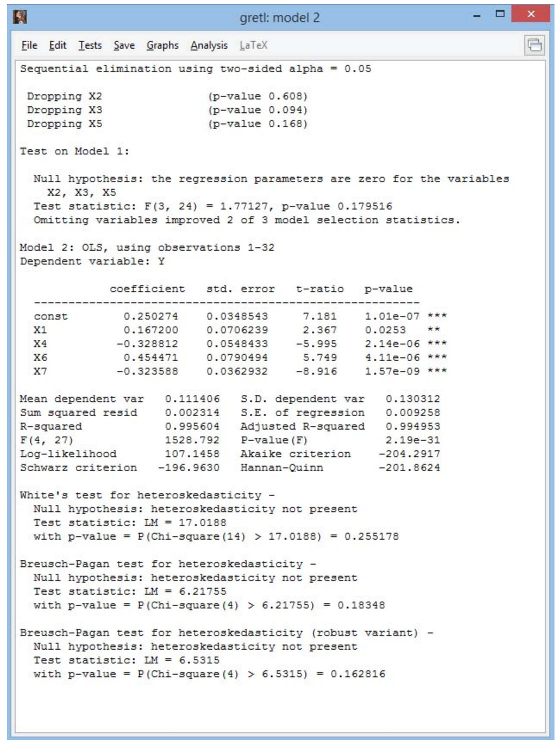
Fig. 3: The test results of model 1 for multicollinearity by inflation factors

$$Y_t = 0.24 + 0.20 \times X_{t1} + 0.01 \times X_{t2} - 0.02X_{t3} - 0.36X_{t4} + 0.12X_{t5} + 0.36X_{t6} - 0.31X_{t7} + \varepsilon_t \quad (1)$$

As seen in Fig. 2, regression has R^2 close to 1, the regression is significant on the whole by Fisher test (p (F)<0.01) and by t-test the regression coefficient by regressor X2-turnover of current assets is not significant that can be the result of collinearity. The sign of the coefficient by regressor X3-the turnover of current assets, does not correspond to the working hypothesis of the model. This situation resulted from regressor multicollinearity. The negative consequences of multicollinearity are inaccurate linear correlation coefficients with the dependent variable-the net return on assets, reduction in the accuracy of regression coefficients, incorrect operation of the Student's test when testing the significance of a coefficient of regressor. To identify multicollinearity (Fig. 3) and redundant collinear factors, calculation of inflation regression test is applicable (Wooldridge, 2009; Hill *et al*., 2012; Garcia *et al*., 2016; Gomez *et al*., 2016):

$$VIF_j = \frac{1}{1 - R_j^2}$$

where, R_j^2-R^2 in the specific regression equation for the jth factor. VIF_j (Variance Inflation Factor) parameter for the jth predictor shows how the estimate of standard deviation for the regression coefficient increases in comparison with the situation if there was no multicollinearity. With VIF_j value of >10, multicollinearity is possible. As seen in Fig. 3, the presence of collinearity is confirmed for regressors X1, net return on sales, X3, current liquidity ratio; X4, the ratio of current liabilities to receivables; X6, the share of payables in the loan capital, %; X7, the ratio of debt capital and assets. It is possible to exclude correlated regressors to eliminate duplication of information. Using Gretl, we eliminate sequentially redundant variables, using two-sided p = 0.05.



gretl: model 2

File Edit Tests Save Graphs Analysis LaTeX

```
Sequential elimination using two-sided alpha = 0.05

 Dropping X2                (p-value 0.608)
 Dropping X3                (p-value 0.094)
 Dropping X5                (p-value 0.168)

Test on Model 1:

  Null hypothesis: the regression parameters are zero for the variables
    X2, X3, X5
  Test statistic: F(3, 24) = 1.77127, p-value 0.179516
  Omitting variables improved 2 of 3 model selection statistics.

Model 2: OLS, using observations 1-32
Dependent variable: Y

             coefficient   std. error   t-ratio   p-value
  --------------------------------------------------------
  const        0.250274    0.0348543     7.181    1.01e-07 ***
  X1           0.167200    0.0706239     2.367    0.0253   **
  X4          -0.328812    0.0548433    -5.995    2.14e-06 ***
  X6           0.454471    0.0790494     5.749    4.11e-06 ***
  X7          -0.323588    0.0362932    -8.916    1.57e-09 ***

Mean dependent var   0.111406   S.D. dependent var   0.130312
Sum squared resid    0.002314   S.E. of regression   0.009258
R-squared            0.995604   Adjusted R-squared   0.994953
F(4, 27)             1528.792   P-value(F)           2.19e-31
Log-likelihood       107.1458   Akaike criterion    -204.2917
Schwarz criterion   -196.9630   Hannan-Quinn        -201.8624

White's test for heteroskedasticity -
  Null hypothesis: heteroskedasticity not present
  Test statistic: LM = 17.0188
  with p-value = P(Chi-square(14) > 17.0188) = 0.255178

Breusch-Pagan test for heteroskedasticity -
  Null hypothesis: heteroskedasticity not present
  Test statistic: LM = 6.21755
  with p-value = P(Chi-square(4) > 6.21755) = 0.18348

Breusch-Pagan test for heteroskedasticity (robust variant) -
  Null hypothesis: heteroskedasticity not present
  Test statistic: LM = 6.5315
  with p-value = P(Chi-square(4) > 6.5315) = 0.162816
```

Fig. 4: OLS estimate of multiple regression of net return on assets after eliminating redundant variables

The presented calculations lead to the conclusion that predictors X2, the turnover of circulating assets; X3, current liquidity ratio; X5, ratio of receivables and payables have to be excluded from Eq. 1. Thus, the exclusion of collinear factors yielded a model of net return on assets (Eq. 2):

$$Y_t = 0.25 + 0.17X_{t1} - 0.33X_{t4} + 0.45X_{t6} - 0.33X_{t7} + \varepsilon_t \quad (2)$$

As seen in Fig. 4, Model 2 has R^2 close to 1, the regression is significant in the whole by Fisher test (p (F)<0.01), all regression coefficients are significant by Student's test. Testing the regression residuals for heteroskedasticity in Model 2 (White's test, Breusch-Pagan test) with a probability of 90% shows the regression residuals homoscedasticity and confirms compliance with the second OLS background. To compare

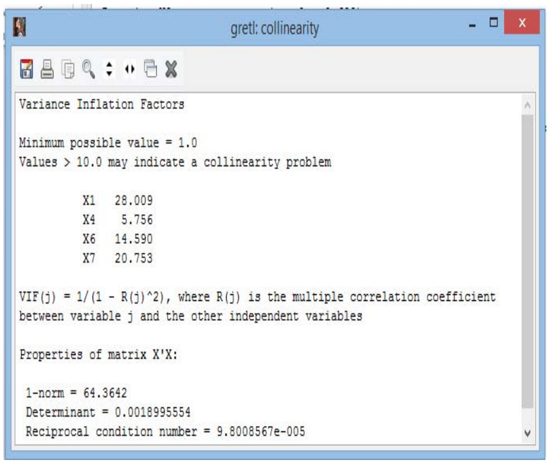


Fig. 5: The results of Model 1 testing to multicollinearity by inflation factors

Models 1 and 2 with the aim of selecting the best one, we will use Schwarz and Hannan-Quinn information criteria, showing goodness of fit, all things being equal. These information criteria introduce penalty function for increasing the number of regressors in the model. Ceteris paribus, the preferred model is the one where the values of these criteria lower. In Model 2, Schwartz and Hannan-Quinn criteria are higher than in Model 1 which does not allow to choose Model 2. Calculation of inflation regression test for Model 2 confirms the presence of collinearity in regressors X1, X6, X7 (Fig. 5).

Therefore, to obtain the best prognostic characteristics, to preserve reliability and information content of modeling, we will do ridge regression that is widely represented in the studies (Garcia *et al*., 2016; Chandrasekhar *et al*., 2016). Ridge regression is a method of down-weighting, the method of selecting factors. It is often used to prevent over-fitting where independent variables are correlated with each other. The consequence is ill-conditioning of the factors matrix and the instability of estimates for regression coefficients. In ridge regression, correlation matrix diagonal adds constant in order to equal all the diagonal elements of the correlation matrix to 1.0 (by = 0 the ridge regression estimators are converted to the least squares estimators). In other words, the ridge regression artificially shrinks the correlation coefficients to calculate more robust estimates of the regression coefficients. If the classic least squares involve a vector-matrix estimation of regression coefficients from the known formula: $B = (X'X)^{-1} \times X'Y$, the addition of parameter solves the problem of ill-conditioning of matrix. Ridge estimators are shifted unlike the OLS estimators. However, it is proved that there exists such λ in which the ridge estimator is more efficient than OLS estimators: $B = (X'X+\lambda I)^{-1} \times X'Y$. Results of ridge regression, received in Gretl for the analyzed enterprise are shown in Table 1.

Table 1: Results of ridge regression

Const.	Regression coefficients	t-statistics ($t_{001;27} = 2.77$)	SE	R^2
a1 = 0.1	a = 0.250	12.66	0.014	0.878
	b1 = 0.320	9.44	-	-
	b4 = -0.120	-3.85	-	-
	b6 = 0.220	5.71	-	-
	b7 = -0.310	-14.23	-	-
a2 = 0.2	a = 0.240	11.79	0.020	0.782
	b1 = 0.310	8.62	-	-
	b4 = -0.090	-2.70	-	-
	b6 = 0.200	5.19	-	-
	b7 = -0.294	-12.61	-	-
a3 = 0.3	a = 0.230	10.54	0.026	0.698
	b1 = 0.290	7.55	-	-
	b4 = -0.080	-2.03	-	-
	b6 = 0.190	4.61	-	-
	b7 = -0.280	-10.81	-	-
a4 = 0.4	a = 0.230	9.52	0.031	0.631
	b1 = 0.270	6.69	-	-
	b4 = -0.070	-1.63	-	-
	b6 = 0.180	4.16	-	-
	b7 = -0.260	-9.41	-	-

On the basis of the highest determination coefficient and the minimum standard error of the model we choose ridge-estimators for λ = 0.1 and write down the final multi-factor linear regression model of the net return on assets:

$$Y_t = 0.25 + 0.32X_{t1} - 0.12X_{t4} + 0.22X_{t6} - 0.31X_{t7} + \varepsilon_t \quad (3)$$

From Model 3 with a probability of 99% it is followed that with a growth of 1% in the net return on sales, net return on assets increases by an average of 0.32% at a constant level of other factors, the growth in the ratio of current liabilities to receivables by 1% increases net return on assets by an average of 0.22%. Empirically observed decrease of the ratio of current liabilities to receivables by one point increases net return on assets by an average 0.12 % when the ratio of debt capital and assets decreases by one point, net return on assets grows by an average of 0.31% in a positive financial leverage.

CONCLUSION

Although, there is a wide variety of approaches to identify and differentiate internal factors and indicators of investment appeal of enterprises, for example such factor as the net return on assets, econometric models have a number of advantages. These advantages include the ability to assess the contribution of each considered factor in the model variation of the net return on assets, to verify the adequacy and relevance of the constructed model, to predict changes in the found dependencies in the future. The performed regression analysis allowed us to confirm the working hypothesis. Since the models were built on real data for a certain period of time, the identified patterns are to some extent, reliable and suitable for further use when analyzing the current situation and identifying trends in the short term. However, it should be remembered that the model is a simplified interpretation of actual market conditions which implies the possibility of excluding a number of factors from consideration which in combination can have a significant impact on the investment appeal. The models obtained during the study can be used as a tool in decision-making by investors.

REFERENCES

Adkins, L.C., 2014. Using Gretl for Principles of Econometrics. 4th Edn., Oklahoma State University Publisher, Oklahoma City, Oklahoma,.

Altman, E.I., 1968. Financial ratios, discriminant analysis and prediction of corporate bankruptcy. J. Finance, 23: 589-609.

Berestetsky, G.L. and E.S. Sapotnitsky, 2004. Facts and forecasts. Cellulose Pap. Cardboard, 7: 10-15.

Chandrasekhar, C.K., H. Bagyalakshmi and M.R. Srinivasan, 2016. Partial ridge regression under multicollinearity. J. Appl. Stat., 43: 2462-2473.

Cokins, G., 2009. Performance Management: Integrating Strategy Execution, Methodologies, Risk and Analytics. John Wiley & Sons, Hoboken, New Jersey, USA., Pages: 271.

Garcia, J., S. Roman and G. Catalina, 2016. Standardization of variables and collinearity diagnostic in ridge regression. Intl. Stat. Rev., 84: 245-266.

Gomez, R.S., J.G. Perez, M.D.M.L. Martin and C.G. Garcia, 2016. Collinearity diagnostic applied in ridge estimation through the variance inflation factor. J. Appl. Stat., 43: 1831-1849.

Hill, R.C., W.E. Griffiths and G.C. Lim, 2012. Principles of Econometrics, International Student Version. 4th Edn., John Wiley and Sons, Singapore, ISBN: 978-0-470-87372-4, Pages: 792.

Kapler, J.K., 2000. Measuring the economic rate of return on assets. Rev. Ind. Organization, 17: 457-463.

Khasanova, A.S., G.M. Kvon, N.M. Yakupova and F.F. Khamidullin, 2015. Assessment of efficiency of capital investment project implementation of resource-saving technology for the real sector of the economy in Tatarstan republic. Mediterr. J. Social Sci., 6: 155-161.

Khavin, D.V., 2004. Investment appeal of industrial objects of the real estate as the factor of effective re-structuring of the enterprises. Mounting Spec. Works Constr., 9: 17-18.

Samerkhanova, A.A. and E.I. Kadochnikova, 2015. Econometric analysis of the mortgage loans dependence on per capita income. Asian Social Sci., 11: 55-59.

Walsh, C., 2011. Key Management Ratios: The 100+ Ratios Every Manager Needs to Know. Pearson Education, London, England, ISBN:9780273761570, Pages: 393.

Wooldridge, J.M., 2009. Introductory Econometrics: A Modern Approach. 4th Edn., South-Western Cengage Learning Publisher, ?Boston, Massachusetts, ISBN:9780324585483, Pages: 865.