

FACTOR MODELS IN INVESTMENTS AND INCOME FORMING PROCESSES

Nikolaj I. Sidnyaev[§]

Bauman Moscow State Technical University

Abstract. In the article it is described a modern theory of securities portfolio forming. It is offered methods helping the investor to choose his optimum portfolio from an infinite number of possible ones the investor is offered to estimate its expected yields and standard deviation together with all covariations between these securities to solve of a matter on inclusion of each considered security in a portfolio. These estimations help the investor to define a curve of effective Markowitz set. In risk-free rate the investor can find a «tangent» portfolio and define position of linear effective set, as well as he can invest into this «tangent» portfolio and make a loan or give a credit on risk-free rate. Therefore the loan or credit sum depends on preferences of the investor concerning a parity of risk and profitability.

Key words: investor, income forming, factor models, sensitivities, factor uploading, portfolio, factor risk, correlation, attributive sensitivities

JEL code: Q550

1. Principles of factor modeling

Factor model is an attempt to capture the key economic forces that systematically affect the market value of all securities (Sharp and others, 2003; William, 1984; Makmin, 2001; Edronova, Mizikovskij, 1995).

When building a factor model it is implicitly assumed that the yields are correlated only due to a shared response to one or more factors identified in the model. As a result the factor model is a powerful tool for investment portfolio management. It can provide with the necessary information to calculate the expected yields, variances and covariance for each security, which is a necessary condition for determining the curve of the effective Markowitz set. It can also be used to characterize the sensitivity of the portfolio to changes in the factors. In factor models it is assumed that the yield securities react to changes in various factors (or indexes) (Edronova, Mizikovskij, 1995; Sidnyaeva, 2006). In the market model it is assumed that there is only one factor - the yield on the market index. However, attempts to estimate accurately the expected yield, variance and covariance of securities with the multifactor model are potentially more useful than with the market model. This is because the effective yields on

[§] Sidn_ni@mail.ru

securities are sensitive not only to changes in the market index but also to economy in general.

In practice, all investors explicitly or implicitly apply factor models. It happens because it is impossible to consider the interconnection of each security with each other security individually, such as the amount of computation when calculating covariance of securities increases with the increase of the number of the analysed securities.

In economic theory, the most important is to evaluate the forecast results (Sidnyaev, 2003; Sidnyaeva, 2001; Sidnyaev, Smirnov, 2009; 2010). More often these results are not subjected to one or a set of the simplest known functions. Therefore, when evaluating results, it is advantageous to use the approximating polynomials, which can describe the response function. Most often Taylor series with a finite number of members are used as the basic function.

$$F(x) = F(a) + \frac{x-a}{1!} F'(a) + \frac{(x-a)^2}{2!} F''(a) + \dots + \frac{(x-a)^n}{n!} F^{(n)}(a).$$

When using the approximating polynomial Taylor in its original form, there arise several problems associated with finding the derivative, such as the function is unknown. Only the number of its values is available. Therefore, as a rule, the Taylor polynomial is replaced by regression equation:

$$r = b_0 + \sum_{i=1}^k b_i x_i + \sum_{i,j=1}^k b_{i,j} x_i x_j + \sum_{i=1}^k b_{i,i} x_i^2 + \sum_{i,j,n=1}^k b_{i,j,n} x_i x_j x_n + \dots$$

where k - the number of columns in the observation matrix. Linear regression model is used in the paper for investment processes studies.

$$r = b_0 + \sum_{i=1}^k b_i x_i$$

In this case there are $k+1$ unknown coefficients, i.e., for each k -factor experiment it will be enough $k+1$ observations. The measurement results of the output value r in k points in the factor space are realization of normally distributed values. Dispersion of implementation at all points in the factor space should be the same, i.e., the variance should not depend on the absolute values of the magnitude and direction of the factor space. Input variables (factors) are independent quantities, which are measured with an infinitely small error relative to the error of the output value.

2. One-factor models of yield securities

The relationship between the forecast rate of GDP growth and equity returns of the investment company can be expressed as the equation:

$$r_t = a + bB_t + e_t, \quad (1)$$

where r_t - yield stocks for the period t ; B_t is the forecasted growth rate of GDP over the period t ; e_t - unique, or specific yield for period t ; b - sensitivity to a forecasted growth rate of GDP; and a - zero factor for GDP. The element changes from period to period and it depends on the forecasted growth rate of GDP (bB_t). Single-factor model can be generalized in the form of the equation for any securities i in time period t :

$$r_{it} = a_i + b_i F_t + e_{it}, \quad (2)$$

where F_t is the forecasted value of the factor in period t , and b_i - sensitivity of securities i to this factor. If the forecasted value of the factor is equal to zero, the yield of such security would be $a_i + e_{it}$. Note that e_{it} is a random variable with zero mean value and standard deviation σ_{ei} . According to the single-factor model, the expected yields of security i can be written in the following form:

$$\bar{r}_i = a_i + b_i \bar{F}, \quad (3)$$

where \bar{F} denotes the expected factor value. This equation can be used to estimate expected yields. For example, if the expected growth rate of GDP is 3%, it is expected for yield security of the company to be equal to 10% ($4\% + 2 \cdot 3\%$). Single-factor model can also show that the variance of any security i is equal to:

$$\sigma_i^2 = b_i^2 \sigma_F^2 + \sigma_{ei}^2, \quad (4)$$

where σ_F^2 - variance of factor F , a σ_{ei}^2 - random error variance. Therefore if the variance of the factor is equal to 3, and the residual variance is 15,2, according to this equation, the variance of the securities is equal to: $\sigma_i^2 = (2^2 \cdot 3) + 15,2 = 27,2$.

In the single-factor model, it is possible to show that the covariance between any two securities i and j is equal to:

$$\sigma_{ij} = b_i b_j \sigma_F^2. \quad (5)$$

Equations (4) and (5) are based on two important assumptions. First, it is assumed to be the lack of correlation between the random error and the factor. This means that the value of the factor does not affect the magnitude of the random error. Secondly, it is assumed to be the lack of correlation between random errors of any two securities. This means that the magnitude of the random error of a security does not affect the magnitude of the random error of any other securities. The assumption is that the yields of all securities react to a single common factor that greatly simplifies the task of determining the "tangent" portfolio. For

determining its composition the investor must evaluate all of the expected yields, variance and covariance. In the single-factor model, it can be done by evaluating a_i , b_i , and σ_{ei} for any of risky securities N. It is also necessary to have the expected value of the factor and its standard deviation. Using these estimates in equations (3), (4) and (5), one can calculate the expected return, variance and covariance of securities. Using these parameters one can define the curve efficient Markowitz set. It can be defined "tangent" portfolio for an actual risk-free rate.

It is known that diversification leads to the averaging of market risk and mitigating own risk. This peculiarity applies to any single-factor model, if one replaces the words "market" and "private" to "factor" and "non-factor". The first term in the right-hand side of equation (4) ($b_i^2 \sigma_F^2$) is called security factor risk, and the second (σ_{ei}^2) is called security non-factor risk.

In the single-factor model, the variance of the portfolio is given by the expression:

$$\sigma_p^2 = b_p^2 \sigma_F^2 + \sigma_{ep}^2, \quad b_p = \sum_{i=1}^N X_i b_i, \quad \sigma_{ep}^2 = \sum_{i=1}^N X_i^2 \sigma_{ei}^2 .$$

(6)

Equation (6) shows that the total risk of any portfolio can be represented in the form of two components, similar to the two components of the total risk of the individual securities shown in equation (4). As the portfolio becomes more diversified (i.e., containing more securities), each security X_i becomes smaller. However, this will not lead to significant increase or decrease of b_p , if it was not made an attempt to do this by adding securities with relatively small or large values b_i respectively. This is due to the fact that b_p is simply the weighted average of security sensitivity b_i in which the weights are values X_i . Therefore diversification leads to averaging of risk factor. However, as portfolio becomes more diversified, it can be expected the reduction of non-factor risk σ_{ep}^2 . Assuming that one and the same amount is invested in each security, this equation can be rewritten with the substitution of X_i to $1/N$ in the following way:

$$\sigma_{ep}^2 = \sum_{i=1}^N \left(\frac{1}{N} \right)^2 \sigma_{ei}^2 = \left(\frac{1}{N} \right) \left[\frac{\sigma_{e1}^2 + \sigma_{e2}^2 + \dots + \sigma_{eN}^2}{N} \right].$$

(7)

The value inside the square brackets is the average non-factor risk for individual securities. But non-factor portfolio risk is only $1/N$ part of this value due to the multiplier $1/N$ in front of the parentheses. As the portfolio becomes more diversified, the number N of securities enlarges in it. Here $1/N$ decreases, that, in turn, reduces non-factor risk of the portfolio, i.e. diversification reduces non-factor risk.

3. Multi-factor models of yields

Unlike single-factor models multi-factor model of yield securities, taking into account these different effects may be more accurate. As an example, let's consider a model in which it is assumed that the process of income generation includes two factors. Two-factor model for period t is written as the following equation:

$$r_{it} = a_i + b_{i1}F_{1t} + b_{i2}F_{2t} + e_{it}, \quad (8)$$

where F_{1t} and F_{2t} - two factors influencing on yields, and b_{i1} and b_{i2} - sensitivity of security i to these two factors. As in case of single-factor model, e_{it} - random error, a_i - expected yields of security i under the condition that each factor value is equal to zero. Observation points are approximated by plane obtained using the statistical method of multiple regressions (Adronov, Mizikovskiy, 1995). The plane for any security is described by equation (8):

$$r_t = a + b_1B_t + b_2I_t + e_t$$

Slope of plane in the direction of the growth rate B_t (b_1) is the sensitivity of the company's securities to changes in growth rate B_t . Slope of plane in the direction of inflation rate I_t (b_2) is sensitivity of the securities to changes in inflation rate.

Under two-factor model it is necessary to estimate four parameters a_i , b_{i1} , b_{i2} and standard deviation of random error designated as σ_{ei} . For each factor it is necessary to estimate two parameters - expected value of each factor (\bar{F}_1 and \bar{F}_2) and dispersion of factor (σ_{F1}^2 and σ_{F2}^2). Finally, it is required to estimate the covariance of factors - $COV(F_1, F_2)$. Using these estimates, the expected yields of any security i can be calculated by the following formula:

$$\bar{r}_i = a_i + b_{i1}\bar{F}_1 + b_{i2}\bar{F}_2, \quad (9)$$

According to the two-factor model dispersion of any security i is equal to:

$$b_i^2 = b_{i1}^2\sigma_{F1}^2 + b_{i2}^2\sigma_{F2}^2 + 2b_{i1}b_{i2}\text{cov}(F_1, F_2) + \sigma_{ei}^2. \quad (10)$$

If in the given example the dispersion of first (σ_{F1}^2) and second (σ_{F2}^2) factors is equal to 3 and 2,9 correspondently, and their covariation [$COV(F_1, F_2)$] is equal to 0,65 than the dispersion for securities of a company will be 32,1 $\left[(2,2^2 \cdot 3) + (-0,7^2 \cdot 2,9) + (2 \cdot 2,2 \cdot (-0,7) \cdot 0,65) + 18,2 \right]$, such as their sensitivity and dispersion of random error are equal to 2,2-0,7 и 18,2

correspondently. Similarly, according to the two-factor model, the covariance between any two securities i and j is defined by the following formula:

$$\sigma_{ij} = b_{i1}b_{j1}\sigma_{F1}^2 + b_{i2}b_{j2}\sigma_{F2}^2 + (b_{i1}b_{j2} + b_{i2}b_{j1})\text{cov}(F_1, F_2).$$

(11)

As in single-factor model, when the expected return, variance and covariance are calculated using the above equations, the investor can use the "optimizer" to obtain the curve of the efficient Markowitz set. As in the single-factor model, the sensitivity of a portfolio to a specific factor in a multi-factor model is equal to weighted average of securities sensitivity, where weights are equal to fractions in which the funds are invested in securities. It can be demonstrated if we remind that portfolio yields are equal to weighted average yields of its securities:

$$r_{pt} = \sum_{i=1}^N X_i r_{it}.$$

(12)

Substitution in the right-side of equation (8) instead of r_{it} gives:

$$\begin{aligned} r_{pt} &= \sum_{i=1}^N X_i (a_i + b_{i1}F_{1t} + b_{i2}F_{2t} + e_{it}) = \left[\sum_{i=1}^N X_i a_i \right] + \left[\sum_{i=1}^N X_i b_{i1} F_{1t} \right] + \left[\sum_{i=1}^N X_i b_{i2} F_{2t} \right] + \left[\sum_{i=1}^N X_i e_{it} \right] = \\ &= a_p + b_{p1}F_{1t} + b_{p2}F_{2t} + e_{pt}, \end{aligned}$$

(13)

where

$$a_p = \left[\sum_{i=1}^N X_i a_i \right]; b_{p1} = \left[\sum_{i=1}^N X_i b_{i1} \right]; b_{p2} = \left[\sum_{i=1}^N X_i b_{i2} \right]; e_{pt} = \left[\sum_{i=1}^N X_i e_{it} \right].$$

It should be noted that the sensitivities of the portfolio b_{p1} and b_{p2} are weighted average of the corresponding individual sensitivities b_{i1} and b_{i2} . In the two-sector factor model the process of income forming on securities has the same general form as in the two-factor model described by equation (9). Now in the two-sector factor models F_1 and F_2 denote sector-factors 1 and 2 respectively. Further, any specific security belongs one of the sectors 1 or 2 but not to both. By definition, the sensitivity of the security to the sector factor to which this security does not belong to is assumed to be zero. This means that either b_{i1} or b_{i2} is equal to zero depending on which of the sectors the security i does not belong to. As an example let's take companies A and B. Two-sector factor model for A is given by the equation:

$$r_A = a_A + b_{A1}F_1 + b_{A2}F_2 + e_A.$$

However, due to the fact that A as an industrial company belongs to sector 1, the coefficient b_{A2} is assigned a value of zero. Then equation (12) reduces to equation: $r_A = a_A + b_{A1}F_1 + e_A$.

So in the two-sector factor model it is necessary to estimate only values a_A , b_{A1} and σ_{eA} for A. By way of illustration, let's see that in the two-factor model it is necessary to estimate the values of a_A , b_{A1} , b_{A2} and σ_{eA} . Similarly, such as B belongs to the non-industrial sector, the two-sector factor model for it is as following

$$r_B = a_B + b_{B1}F_1 + b_{B2}F_2 + e_B, \quad (14)$$

that simplifies to equation:

$$r_B = a_B + b_{B2}F_2 + e_B, \quad (15)$$

such as b_{B1} is assigned zero value. As a result in the two-sector factor model it is required to estimate only values a_B , b_{B2} and σ_{eB} . In general, while in the two-factor model for each security it is to be estimated four parameters (a_i, b_{i1}, b_{i2} and S_{ei}) in the two-sector factor model it should be estimated only three parameters (a_i, σ_{ei} and either b_{i1} , or b_{i2}). With these estimates, together with estimates for F_1 , F_2 , σ_{F1} and σ_{F2} , the investor can apply equations (9) and (10) to calculate the expected returns and variances for each security. Paired covariances can be estimated using equation (9). This will enable the investor to determine the curve of effective Markowitz set, then the tangent portfolio for an actual risk-free rate.

For the case in k factors a multifactor model can be written in the following form:

$$r_{it} = a_i + b_{i1}F_{1t} + b_{i2}F_{2t} + \dots + b_{ik}F_{kt} + e_{it}, \quad (16)$$

where each security is characterized by k sensitivities, one for each of the k factors.

Equation (16) can contain both factors and sector-factors. For example, F_1 and F_2 can denote GDP and inflation, while F_3 and F_4 can be industrial and non-industrial securities yields, respectively. Consequently, the securities of each type would have three sensitivities: b_{i1}, b_{i2}, b_{i3} for industrial and b_{i1}, b_{i2} and b_{i4} non-industrial securities. As an equation the factor model for the securities looks as follows:

$$r_{it} - r_{ft} = a_i + b_{i1}(r_{Mt} - r_{ft}) + b_{i2}S_t + b_{i3}H_t + e_{it}. \quad (17)$$

The first factor ($r_{Mt} - r_{ft}$) is equal to the difference between the monthly returns of the security market and one-month treasury bills. The factor of capitalization (St) can be thought of as the difference in monthly returns on the two indices - the indices of the "small" and "large" securities. Relationship factor of book value to market (H_t) is also the difference in monthly returns on the two indices - the indices of the securities of high and low ratio of book value to market one. Index of securities with a high ratio consists of those securities that are in the upper third of the values range H_t , and index of securities with low ratio consists of those securities that are in the bottom third. It can be identified two factors that explain the monthly yield on the securities. In the equation their factor model for securities looks as follows:

$$r_{it} - r_{ft} = a_i + b_{i1}T_t + b_{i2}D_t + e_{it}.$$

(18)

4. Method of spatial sampling

Method spatial sampling is completely different from the method of time series. In the latter method, the factor values are known, and their sensitivities are evaluated. Then the analysis is made for the same securities on a large number of time intervals, and then for another security, etc. In the method of spatial sampling the sensitivities are known and factor values are estimated. In this method, the sensitivities are sometimes named attributive. The analysis by this method is made within one time interval for a group of securities, then within another time interval for the same group of securities, etc. To illustrate the method of spatial sampling let's see the examples of one-factor and two-factor models.

In some periods of time securities with low market capitalization are generally of higher returns than securities with high market capitalization. In other periods quite the opposite can be true. There are many models in the method of spatial sampling using size index that is often defined as the logarithm of the total market value of securities placed by the company. The slope of the straight regression line gives estimation to the rate of capitalization in an actual period of time. However, this procedure has some drawbacks. Securities with large capitalization usually have a higher rate of dividends. Therefore, the difference in yields between securities with high and low capitalization may be partially related not only to differences in capitalization but with a difference in rate of dividend. The estimated rate of capitalization may partially reflect the influence of the true rate of dividend. The problem is symmetric in the sense that the estimated rate factor of the dividend may also partially account for the actual capitalization. This problem can be partially resolved by comparing the yield with the amount of capitalization and rate of dividend at the same time and using the statistical method of multiple regressions. Each security is a point on the three-dimensional diagram. Approximation of data using plane is normally estimated by the method of multiple regressions using the regression equation:

$$r_{it} = 7 + 0,4b_{i1t} - 0,3b_{i2t} + e_{it} ,$$

(19)

where b_{i1t} and b_{i2t} denote the rate of dividend and capitalization of securities i for period of time t , correspondently. In general the regression equation for two-factor model is written in the following way:

$$r_{it} = a_t + b_{i1t}F_{1t} + b_{i2t}F_{2t} + e_{it} ,$$

(20)

where a_t denotes zero factor i in period of time t , and two factors are denoted as F_{1t} and F_{2t} .

Conclusions and recommendations

Factor model corresponds to the process of income forming, linking security yield with changes in one or more common factors. Any aspect of security yield being unexplained by the factor model is specific for the actual security, and, it is consequently, uncorrelated with the specific components of the yields of other securities. The market model is a particular example of factor model in which the factor is the return on the market index. Yield securities react to common factors that greatly simplify the task of the curve calculation of effective Markowitz set. The sensitivity of a portfolio to a factor is equal to the weighted average of the sensitivity for its constituent securities. Thereby the weights are the fractions in which the securities are included in the portfolio. Full risk for securities consists of the risk factor and non-factor risk. Diversification leads to the averaging the factor risk or reduces non-factor risk.

Bibliography

1. Edronova, V. N., Mizikovskij, E.A. (1995). Account and Analysis of Financial Actives: Actions, Bonds, Bills. - Moscow: *Finance and statistics*, p. 272.
2. Sharp, U., Alexander, G., Bailey, Dzh. (2003). *Investments*: translated from English-Moscow: INFA-M, 2003.-XII, 1028 p.
3. Karlin, T.R., Makmin, A.R. (2001). *Financial Reports Analisys (on the basis of GAAP)*; Textbook –Moscow. p. 448.
4. Sidnyaev, N.I. (2003). On Self-return Process of Time Series at Modeling of the Investment Projects. Theses of Reports of Scientific-practical Conference «Methods on Teaching the Mathematical Cycle for Economic Specialties». – Ulyanovsk. p.57-58.
5. Sidnyaeva, O. N. (2006). Methodological Bases on use of Cumulative Sums at Monitoring Systems Development in Administrative Decision-Making. *Methodology in Management for Highly Technological Enterprises (Materials of Inter-regional Scientifically-practical Conference-meeting, on May, 31st 2006, Moscow*. p.41-43.
6. Sidnyaeva, O. N. (2001). Institutional Factors of Economic Growth: Fiscal Federalism. *Transformation Processes and Economic Growth in Modern Russia: Collection of scientific articles. Release the first*. Moscow. p. 184.

7. Sidnyaev, N.I., Smirnov, A.V. (2009). Risk Analysis and Management. Forecast of Catastrophic Events. *Materials on XVI International Conference on Computing Mechanics and Modern Applied Program Systems (VMSPPS' 2009), May, 25-31th, 2009, Alushta.* – Moscow. pp.646-649.
8. Sidnyaev, N.I., Smirnov, A.V. (2010). Optimization Theory for Portfolio Financial Risks. *Collection of Reports of All-Russia Scientific-practical Conference «Modern Problems in Innovative Economy of Regions in Russia: Theoretical and Practical Aspects», on April, 23-24rd, 2010, branch GOU VPO «Dagestan state technical university».* Derbent, pp. 328-331.
9. William, F. (1982). Sharpe, «Factors in New York Stock Exchange Security Returns, 1931-1979», *Journal of Portfolio Management*, 8, no. 4 (Summer 1982), pp. 5-19.