

# Chapter 1

## Hedging

### 1.1 Introduction to Hedging

Human life as well as any human activities are inextricably associated with an element of uncertainty. The assumption of uncertainty implies the concept of risk that is part of social system in various forms of understanding. The risk is generally related with some negative meaning. In the field of finance it is most associated with some diversion from an expected state, more precisely the risk is embodied in deviation of expected return. It is not necessary, that the divergence from expected state must not be solely a financial loss. The uncertainty could be demonstrated equally by the positive deviation with higher profit than expected. So it is not only limited to financial suffering. Although, this could be the major threat of the market attendance.

The risk is commonly divided into two components in the financial system. One part of the whole financial risk is a unique risk. Sometimes called unsystematic, undiversified, residual or idiosyncratic risk [[Beja1972](#)]. The concept of financial risk is applied when financial sources are placed in the financial market no matter if in the form of investment or speculation. In theory, but in practical application as well, it is common to work with multiple assets concept. A reduction of residual risk is feasible applying appropriate algorithm for assets allocation (diversification). Apparently, a complete elimination of this risk could be achieved [[Elton1997](#)].

The second part of the overall financial risk is referred as systematic part of risk. This category of risk is also called as non-diversifiable risk. All actors in the financial market face the systematic risk, because every asset is exposed to market risk [[Frenkel2005](#)]. Its existence can only be accepted not removed. What applies to the financial markets could be apply without reservations to the commodity market as well. As noted by [[Garner2010](#)]: *"Producers and users of commodities are constantly faced with price and production risk due to an unlimited number of unpredictable factors including weather, currency exchange rates, and economic cycles"*. The embodiment of the risk is actually the uncertainty arising with the nature of the markets. It is inherent to entire financial market. However, that does not mean that active trading subjects on the market could not influence the impact of the systematic risk. Hedging is a financial operation, that aims to reduce the impact of non-diversible risk [[Collins1999](#)]. On the other hand, some authors pointed out that only part of the systematic risk could be eliminated [[Kolb2014](#)]. Actually it is a closure of positions hold in assets. Growth or decline in the price of one asset is offset by the opposite movement of hedging assets price.

A long and short hedge position could be distinguished from the perspective of trade. The difference depends on whether the asset exposed to the systematic risk intends to be bought, or the asset is already owned and later on will be sold. If it is the right purpose, then an opposite operation, i.e. sell of hedging assets on financial market, must be done. Therefore,

such operation is called short hedging [[Rutledge1977](#)]. Other kind of situation in trade relations constitutes such activities where the short sell will be realized.

1

Short sell is executed on behalf of securities loan. A similar strategy will be applied, when a certain asset should be bought in the future. For instance an essential asset for an entrepreneurial activity. Thus, in order to guard against an eventual rise in the price of intended asset a sell of high correlated asset to the price of considered asset is required. The rising price of the asset bought at present will offset an undesirable price growth of asset purchased in the futures. The above strategy is obviously long hedging [[Bessembinder1992](#)].

Derivatives are widely used for hedging purposes [[Bingham2013](#)]. Thus, the underlying of derivative is the object of hedging [[Poitras2002](#)]. The class of standard derivatives, like forward, future, swap and future were commonly used [[Hull2006](#)]. Development of financial engineering has caused usage of more sophisticated and complex instruments in recent years [[Avellaneda1995](#)].

2

However, a diversion of their primary function could be observed in the last two decades [[Bartram2009](#)]. A growing interest in these instruments can paradoxically lead to increase of market risk, since a large part of the transaction volume is driven by speculative intention [[Stiglitz2000](#)].

3

## 1.2 Literature review

An interest in hedging can be traced in scientific circles in the first half of the 20's century. The pioneers who contributed to the research in this field predominantly focused on usage of future contract in agricultural products [[Howell1938](#)]. Scientists have noted the potential that a standardized derivative instrument can offer due to price fluctuations. [[Yamey1951](#)] states: "*The practice of hedging, by buying and selling futures contracts on organized produce exchanges, enables manufacturers and merchants to cover themselves against adverse movements of prices of raw materials in which they deal*". Likewise, he points out that such a protection against the risk may not be sufficient, since hedging may not be perfect. In principle, hedging was regarded as usage of futures which should guarantee a protection against undesirable price movement. The methodology of protection was based on closure of opposite trade positions in the ratio 1:1, namely the amount of spot volume should be protected with same amount of futures. Arguments for such manner was advocated by very similar price behavior in both considered assets [[Grafl1953](#)]. Thus, a potential loss on spot price could be eliminated by the gain from futures. Considering prices of underlying and futures were determined by identical factors, hedging with same weight appeared to be appropriate [[Howell1948](#)].

The fifties were revelatory for finance. The key achievement was the birth of modern portfolio theory [[Markowitz1952](#)]. New optimization technique could be employed to find an appropriate hedge ration thanks to the contribution of Harry Markowitz [[Telser1955](#)]. Since this point, the attention of scientists did not restrict merely on statistical characteristics of separate assets, but also the mutual interaction among them. This handling has enabled to improve the benefit arising from affinity within return and risk. Said circumstance motivated further evolution in hedging research. The trivial procedure of same weights could be put aside owing to application of utility function made known in modern portfolio theory.

One of a cardinal contribution in scientific literature was the article Hedging Reconsidered [[Working1953](#)]. The author pointed out three major economic effects of hedging. First, the

risk reduction cause fewer bankruptcies of companies with positive effects on society and the whole economy. The forthcoming level of spot prices could be estimated more accurate. He has mentioned the positive impact on commodity stocks as well. The radical thought heretofore was a revolutionary look at the role of hedging. [Working1953a] highlights an incorrect understanding of hedging. He states that protection against potential financial loss is a secondary function of hedging. The primary function according his opinion is usage for arbitrage purposes [Working1953a]. The scientific literature refers to this funding in later years too [Ederington1979], [Cicchetti1981], [Garbade1983], [Tomek1987].

[Grafl1953]discussed also the perception of insufficient ability of hedging to protect against price risk in the fifties. Author has examined the ability of futures to provide risk reduction of potential losses. He has considered the concept of hedging effectiveness and has analyzed the degree of efficiency. He has expressed his opinion about changeability in hedging effectiveness. In consonance with empirical research hedging is showing dynamic development. The date for his study was examined from Chicago Mercantile Exchange. Futures with near or second near month was used for hedging scope and the researched commodities were corn, wheat and oats.

Subsequently, the scientific improvement in the hedging issue has brought [Johnson1960]. His contribution is particularly in the area of quantification. In his view, the price and subsequently return and price risk is represented as a process of random variable [Johnson1960]. Thus, the price risk is identified as the variance of price change over time. The author has defined the optimization process to determine the weights of hedge instrument represented by futures. The knowledge elaborated in modern portfolio theory was adopted in his consideration. Weights of futures are calculated solving optimization problem in utility function. The objective function was the portfolio variance. Breakthrough is that he can clearly identify a portfolio risk depending on the linear relation between spot and futures. Furthermore, Johnson additionally has developed a methodology for measuring the effectiveness of implemented hedging. The inputs are the variance of unhedged asset (spot) and hedged portfolio (combination of spot and futures). The hedge effectiveness is referring about percentage decrease in the variance. What is certainly revealing, is the importance of strong correlation in both assets for effective hedging [Johnson1960].

Stein has adopted the concept of expected return to the hedging problem. The idea follows difference of current spot and futures price with the expected state of both assets. In his reasoning the carrying costs are reflected as well. Similarly to Johnson is considering the portfolio variance as risk of hedged position. In addition, he has discussed the graphical interpretation of dependence between expected return and risk. Furthermore, Stein has worked with the theory of convex indifference curves [Stein1961]. He argues its shape by declining income utility referring to [Tobin1958].

Scientific papers in the successive years drew from the learning of modern portfolio theory simultaneously using futures. In some studies the subject of interest is the measurement risk reduction on ex post data. Among the known authors in this area belongs [Ederington1979]. He is examining the futures weights by the ordinary least squares model on empirical data. The regression is done on percentage changes of prices.

<sup>4</sup>

[Ederington1979] introduced the term basis risk and claims: *"A hedge is viewed as perfect if the change in the basis is zero"*. The measure of hedging effectiveness is describe by the coefficient of determination.

<sup>5</sup>

[[Cicchetti1981](#)] has examined the hedging effectiveness on the money market. His study is focused on the treasury bills traded in Chicago Mercantile Exchange. In his work he was actively referring to Ederington in his paper. [[Dale1981](#)] investigated the hedging effectiveness on the foreign currency market. He adverted to the work of Working. He is researching in addition the market demand and supply. Similarly oriented work represented [[Hill1981](#)]. Examining the same underlying asset was provided by [[Hsin1994](#)]. But moreover, an option is considered as hedge instrument too. Reverting to agricultural commodities was the paper provided by [[Wilson1982](#)]. The subject of their measure was wheat. [[Cotter2006](#)] introduced a modern view to hedging effectiveness. He reminded some lack of the standard measurement and simultaneously shown that different measure could provide different results. He suggested to use the concept of Value at Risk in hedging effectiveness.

Other scientific area in hedging exploring is testing the stability of hedge ratio. This field investigated [[Grammatikos1983](#)]. The authors reminded on the characteristics of previous research. The data processing could be a shortcoming, since the analyses were based on large data period. However, it could be pitfalls. They focused in their analysis on international money market. Specifically, the research analyzed Swiss franc, Canadian dollar, British pound, German mark and Japanese Yen. The results indicated the unsuitability of the stable hedge ratio hypothesis.

Similar results were confirmed by other studies [[Grammatikos1986](#)], [[Eaker1987](#)]. [[Malliaris1991](#)] asked a question if more input data for the analysis could provide better results, because such a dataset will include more information. In contrast, he puts a hypothesis of instability of beta coefficient. Thereby the first assumption was refused. The reason was that not all information incorporated in the processing dataset are significantly relevant for hedging purposes. In other words, the data from remote history did not provide much information for current data. His research confirmed the hypothesis that the hedge ratio showed instability over time. Simultaneously added that the beta coefficient were not significantly different from one. And finally he persuaded that foreign currency futures are convenient tools for hedging.

Another popular hedging area within the scientists was comparing the performance of optimum hedge ratio with the pay off produced by naive portfolio. Among such studies belong the work of [[Eaker1987](#)]. They compared multiple methods of hedging with the naive portfolio. Similarly oriented work is presented by [[Hammer1988](#)]. Three methods of hedge optimization are compared with naive portfolio in [[Park1995](#)]. The data from S&P 500 and TSE 35 was examined for their analysis. Further paper comparing different form of hedging with naive portfolio was done by [[Bystrom:2003](#)]. He annualized the electricity market Nord Pool. However, how noted by [[Collins2000](#)] it is not always possible to confirm a benefit of "*sophisticated*" and complex econometrics models over performance of naive portfolio.

The three above introduced area of research developed more or less separately. Nevertheless, they are closely related. So [[Marmer1986](#)] decided to examine all three presented area together. The object of his investigation was Canadian dollar and exchange rate futures. The results of Marmer's analysis spoke in favor to optimized approach over naive portfolio. He also rejected the hypothesis of stable hedge ratio. Further, he declared that with raising duration rose the hedging effectiveness as well. Similarly the variability in hedge ratio and risk reduction over time was confirmed by [[Benet1992](#)].

The fundamental shortcomings of previous models was the unsustainable assumptions about stationary of data. It was a matter of time when scientist begin to deal with above given

circumstances. One way how to solve the problem with non-stationarity may be models ARCH, GARCH models or co-integration.

A modern perspective in this issue gave [[Cecchetti1988](#)]. He used Autoregressive conditional heteroscedasticity model ARCH for solve hedge ratio. After successful application of ARCH in the area of financial asset valuation and when the Generalized autoregressive conditional heteroscedasticity model (GARCH) was introduced it is also utilized in hedging [[Engle1986](#)], [[Bollerslev1987](#)]. Among scientists working with GARCH in the field of hedging belong [[Myers1991](#)], who focused on commodity hedging. Six commodities were investigated in his analysis under conditional variance and covariance. [[Ghosh1993](#)] estimated optimal futures hedge ratio for non-stationary data and incorporated long-run equilibrium with short-run dynamics. The underlying asset was S&P 500. He applied Error Correction Model (ECM) in his research. The results of ECM was better to traditional approach. A similar procedure also chosen [[Chou1997](#)]. He was handling the hedging on the Japanese Nikkei Stock Average. Again the results were in favor for ECM to the conventional models.

Further similar study is provided by [[Ghosh1996](#)]. This time the index CAC40, FTSE 100, DAX and NIKKEI were explored. Again the co integration was used and the results confirmed the hedging effectiveness of ECM over standard approach. [[Alexander1999](#)] stated: "*If spread of spot and futures are mean reverting, prices are co-integrated*". In addition, he demonstrated the occasion for using co integration for different purposes like arbitrage, year-curve modeling and hedging. He verified the hedging on European, Asian and Far East Countries.

One of such paper was provided by [[Baillie1991](#)]. They examined the hedging on six commodities. They shown and emphasized how important is to take in consideration the non-stationary in examined data. [[Moschini2002](#)] introduced a new multivariate GARCH parametrization. The authors tested hedging effectiveness on models with time-varying volatility. The better hedging performance was confirmed by [[Yang2005](#)] in multivariate GARCH over classical OLS as well. The examined data was from Australian financial market. [[Lee2007](#)] used Markov regime switching GARCH for estimating minimum variance hedge ratio under time-varying variance. The authors used BEKK-GARCH for the analysis. Their decided to use a complex algorithm because of the changing joint distribution of spot and futures over time. The analyzed area were prices of corn and nickel. They confirmed a better hedging effectiveness by surveyed commodities after using the GARCH model, but at the same time they added that the difference in comparison to other models is not significant.

Other technique for optimization treats the knowledge of portfolio theory likewise. Although, the utility function is adopted from portfolio, it is different to the minimum variance. Since, to the previous optimization, now the extreme value represents the maximum of examined function. Correspondingly to minimum risk optimization the portfolio variance is employed but besides an expected excess return of the hedge instrument.

[[Howard1984](#)] recommended using of Sharpe ratio for estimating hedge ratio.

Unless, there will be an assumption that the expected return of futures is zero, then even this optimization generates identical hedge ratio alike the minimum variance.

Nevertheless, [[Chen2013](#)] warned that Sharpe ratio is not a linear function, which could be problematic. As a consequence solving the optimum, could let to find the minimum instead of the maximum value.

Another optimization technique finding maximum of an objective function is optimum mean-variance hedge ratio [[Hsin1994](#)]. The objective function does not include only portfolio expected return and variance, but also an attitude for risk aversion. With this concept came already [[Heifner1966](#)].

However, he reminded a subjectivity of this parameter. Under certain conditions the hedge ratio derived from optimum mean-variance utility function will be identical to the ratio from minimum variance.

<sup>10</sup>

The concept of optimum mean-variance was later applied by other authors [[Hsin1994](#)], [[Moschini2002](#)], [[Casillo2004](#)].

[[Cotter2010](#)] examined time-varying risk aversion from the observed risk preferences in market participants. According to the risk aversion a long and short hedging was implemented. The authors argued that respecting the concept of time-varying risk aversion outperformed the standard OLS approach. Subsequently, those authors focused on application of different utility functions for risk preferences in hedging. Namely, they employed logarithmic, exponential and quadratic form of utility function for determining risk aversion. The results of their study confirmed significant differences in the optimal hedge ratio upon distinct utility function [[Cotter2012](#)]. The above mentioned authors also emphasized the relevance of asymmetry in the return distribution which could have impact on the hedging effectiveness. They accentuated a shortcoming of minimum variance hedge ratio [[Cotter2012b](#)].

There are frequently referred [[Kolb1992](#)] in relation to Gini application in hedging. A modified version of Gini ratio so called Extended mean Gini coefficient was introduced in their article. The improvement of Gini was done by incorporating an element of risk aversion to the ratio. Additionally, [[Lien1993](#)] emphasized that the parameter of risk aversion plays a crucial role in the estimation. [[Shalit1995](#)] noted a particular problem by comparing the hedging effectiveness between mean-variance and mean-extended Gini coefficient. [[Lien2002](#)] then discussed an economical implementation of the conventional approach for hedging purpose derived from the expected utility maximization paradigm with the new econometrics procedures. Similar [[Chen2013](#)] compared the different methodologies for calculating an optimal hedge ratio. Nonetheless, they asserted that not always a "modern" approach could over-perform the classical OLS.

Emerging studies from the recent years have used more and more sophisticated and complex mathematical tools. A methodology dealing with joint distribution probability function could be one of the examples. So new papers appeared with the models from the copula family to find a proper hedge ratio [[Cherubini2004](#)]. Alternatively, to make things more interesting a combination of copula calculus with other econometrics methods were applied [[Hsu2008](#)], [[Lai2009](#)], [[Lee2009](#)]. Of course, the wavelet transformation could be added to the innovative form of solving the hedging problem as well. [[In2006](#)] tried to solve the problem with time-varying covariance within spot and futures using wavelet. Similarly, [[Fernandez2008](#)] is estimating hedge ratio after wavelet analysis. Wavelet and ECM applied [[Lien2007](#)] too with distinctive results. In his article he came to the conclusion that the time horizon for hedging seems to be important. With growing time horizon of hedged asset the wavelet approach delivered better performance.

Hedging in energy commodities are mostly applied in oil, natural gas and electricity. Typically a high level of volatility is common for all three markets. [[Chen1987](#)] examined the oil hedging. He explained the instability in oil prices with restructuring in oil industry in eighties. But how shown the later oil price process, the high instability in price of oil is

inherent. Nevertheless, he noted that the risk exposure affected primary producers and users and is advocating application of futures contracts as an appropriate tool helping protect against price risk. [Duffie1999] gave a guidance on how to handle with empirical behavior of volatility in the energy sector. He analyzed the problem of stochastic volatility. Simultaneously they described various form of Markovian stochastic volatility models. [Haigh2002] examined the market of crude oil, heating oil and gasoline to reduce the price volatility. The applied model was constructed as a crack spread. They also focused attention on time to maturity effect in futures. The result of their investigation confirmed usage of Multivariate GARCH methodology as beneficial. [Dahlgren2003] specialized in protection against price risk in the power market. Their efforts were focused on the promotion of risk assessment and application of hedging in power sector. [Woo2006] examined the degree of co-integration of natural gas market in the USA. The data for analysis was from Californian market. They recommended to use futures contract traded in New York Mercantile Exchange with underlying spot of Henry Hub to reduce price risk. [Alizadeh2008] focused on the same market in New York but to examined commodity was oil. The authors applied the concept of dynamic hedging and operate with high and low volatility regime. According their results a dynamic hedging provides significant reduction in portfolio risk. [Chang2011a] examined two global oil markets Brent and WTI. In the analysis was used the model with multi-variation volatility models. The results implied different hedge ration in accordance with the used methodology and similarly variable level of hedging effectiveness. They also shown differences in hedging between the two markets [Chang2011a].

## 1.3 Hedge ratio

The first prerequisite for application any hedging strategy is a selection of convenient asset, that should be hedged. A convenient asset must exist for hedging purpose. Convenient assets for hedging are those with high price affinity. Its existence can be simply confirmed by the value of Pearson product-moment correlation coefficient [Calmorin2004]. The absolute value of correlation coefficient close to one is a prerequisite for a successful hedging. Which means that prices of surveyed assets may evince identical or reverse movement. It is irrelevant. The crucial is that the price extent in both assets is identical. When the prices will have a high negative correlation, then the hedging will be executed by realization of the same position, i.e. both assets will be bought or sold. However, the assets show high positive correlation and such situation is closer to real data, then the hedging must be executed by opposite transaction in assets. The long asset has to be hedged by selling other asset and vice versa.

However, in the real world it is rather sporadic to trace perfectly correlated prices over long term. How noted by [Working1953] the circumstance of releasing price tightness between spot futures prices results an ineffective hedging.

In order to apply hedging it is require to determine the weights of asset commonly named hedge ratio. The hedge ratio is referring how much hedging assets should be considered to the exposure.

At the very beginning the hedging was based on the assumption of identical price movement in considered assets. The protection against unintentional price loss should be provided by closing the open position with hedge asset. In other words, the volume of owned assets was matched by the same quantity of hedge assets. Given hedging strategy is also called naive portfolio [DeMiguel2009]. Nonetheless, such protection has been frequently associated with imperfections [Brooks2002]. The inadequate protection was caused by a lack of perfect correlation. Subsequently, an unequal hedge ration began to be applied. Handling a new hedging methodology was enabled due to usage of modern portfolio funding. Afterwards, the

hedge position has been determined by using the optimization of minimum risk. The risk was given in the form of variance. The objective function for solving the extreme value was exactly variance of two assets (variance of portfolio). Where in addition to the solely variance also statistical relation in the form of covariance is considered. Referring to the work of [Markowitz1952] the diversification effect between two assets in order to minimize risk was applied.

The portfolio variance:

$$\sigma_p^2 = w^T \Sigma w = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \sigma_{i,j}$$

Then the quadratic utility function for solving an optimal hedge ratio represented by futures is:

$$\sigma_p^2 = w_s^2 \sigma_s^2 + w_f^2 \sigma_f^2 + 2 w_s w_f \sigma_{s,f}$$

Where  $w_s$  is weight of spot,  $\sigma_s^2$  is variance of spot,  $w_f$  is weight of futures,  $\sigma_f^2$  is variance of futures, and  $\sigma_{s,f}$  is covariance between spot and futures. The weights restriction is distinct to classical optimization in [Markowitz1952]. It is assumed one unit of spot. The object of hedging here is to find an appropriate weight proportion of futures to the spot. Solving the extreme of the objective function will provide following expression:

$$\frac{\partial \sigma_p^2}{\partial w_f} = 2 w_f \sigma_f^2 + 2 \sigma_{s,f} \quad (1)$$

When  $w_f w_s = h^*$  then the optimal hedge ratio is derived as:

$$h^* = - \frac{\sigma_{s,f}}{\sigma_f^2} \quad (2)$$

The sign of the hedge ratio refers about short position in futures [2](#). Testing second order condition we can confirmed that the extreme value of the objective function is minimum [3](#):

$$\frac{\partial^2 \sigma_p^2}{\partial w_f^2} = 2 \sigma_f^2 \quad (3)$$

The ratio is called minimum-variance hedge ratio (MVHR) precisely because of the objective function. In the practical part will be applied the MVHR. Nevertheless, the statistical characteristics and their interaction are going to be estimated according to distinct methodologies. Overall, there are seven distinct techniques to provide the hedge ratio in the analytical part of the dissertation. It concerns the following methods:

- Ordinary least square
- Naive portfolio
- Error Correction Model
- Arch/Garch Model
- Wavelet
- Copula
- Extended Mean Gini Coefficient

## 1.4 Measure of hedge effectiveness

The results of the weights for futures produced by each of the seven above introduced models will be used for hedging the spot prices of the three examined commodities. The obtained hedge ratios are applied on the real data in the following twelve months. Thus, the ability to reduce risk in every method is measured and compared during the specified period.



## Conventional measurement

As soon as the hedge ratio is calculated, the value is applied on the real data to calculate variances and covariance. Evaluation of hedging performance is based on percentage reduction in spot variance compared to portfolio variance. The metrics was made in accordance with the methodology of [[Johnson1960](#)]:

$$HE = \frac{\sigma_U^2 - \sigma_H^2}{\sigma_U^2}$$

HE stands for hedging effectiveness and  $\sigma_H^2$  is variance of hedged portfolio or spot with futures together and  $\sigma_U^2$  is the variance of unhedged portfolio or the variance of spot. It is apparent that the better the futures will match the spot, the lower risk will the portfolio evinces. In other words the risk reduction will be higher and the coefficient HE will be closer to 1 which actually would be the 100 % reduction of risk. On the contrary the closer the value of HE is to zero, the larger imperfection of the hedging is present.

In the calculation of the portfolio variance it must be taken account of the negative futures weights [2](#). The reverse weights leads to the reduction of the covariance risk, since the both variances are additive [4](#):

$$\sigma_p^2 = \sigma_s^2 + h^2 \sigma_f^2 - 2h \sigma_{s,f}; \text{ variance risk } -2h \sigma_{s,f}; \text{ covariance risk .}$$

(4)

It is evident that the closer price co-movement exists, the better hedging will be provided. So the hedging efficiency is determined by correlation, because it is obviously true:

$$\rho = \frac{\sigma_{s,f}}{\sigma_s \sigma_f}$$

The effect of risk reduction decrease unless the correlation release. In the case of significant drop in correlation or the correlation even changes to negative value, then the risk of hedged portfolio paradoxically increase.

[[Cotter2006](#)] highlights deficiencies arising with the usage of very popular measure of hedging effectiveness. That's why he gives a measurement of hedging effectiveness based on the concept Value at Risk [[Cotter2012b](#)]. Then the expression for measure take following form:

$$HE = 1 - \frac{VaR_H}{VaR_U}$$

where VaR corresponds to the (100-x)th percentile of the portfolio over next N days. He applied x=1 and N=1 and the subscripts H,U mark hedged and unhedged portfolio.

<sup>13</sup>

## Alternative metrics

The results of hedging effectiveness can be compared partially by months. However, the question is how to compare the aggregated results over the whole exterminated period. Since it would be problematic to provide test of statistical significance from achieved value in HE between all methods over the intended period.

<sup>14</sup>

Therefore it was advisable to find another measurement to evaluate the appropriateness of each method for examined commodity. A simple comparison provide tools of descriptive statistics such as the mean or the median.

An another option for a comprehensive comparison is the sum of differences between a reference value and the achieved HE. If the reference value is established as one, then the

difference will referring about the residual variance risk in the portfolio. Thus, the relation could be expressed as follow:  $R_r = 1 - \exp(-\lambda_i)$ ,  $\lambda_i = \ln(H/E_i)$ . (5)

Where  $R_r$  is the residual risk and  $\lambda$  is natural logarithm of achieved hedging effectiveness by each model. In fact, the above relation indicates remaining percentage from perfect hedging expressed in variance. Obviously it is possible to derive the risk reduction in absolute value:  $\Delta\sigma = \sigma_U \exp(-\lambda)$ ,  $\Delta\sigma = \sigma_U - \sigma_H$ .

In the equation the terms  $\sigma_U$  and  $\sigma_H$  are the standard deviation of the logarithmic returns of spot and futures prices and here the risk reduction is calculated from standard deviations. Then the standard deviation of the hedged portfolio is possible to write as:  $\sigma_H = \sigma_U (1 - \exp(-\lambda))$ . (6)

Thence, the cumulative residual risk may be used to compare different methods among themselves.

An alternative way for comparing the models is establish a ranking according to their performance in the particular months. The achieved score in all months provides a view about performance under particular measurement. However, this evaluation do not take in the account the overall effect of risk reduction over the whole period.

## Chapter 2

# Applied models for determining the MVHR

## 2.1 Ordinary least square

Let  $X$  be a matrix with dimension  $n \times 2$ . The matrix will contain one independent variable and a constant term in the first column. In the system will be  $n$  observations (rows). The matrix  $X$  is regarded as independent variable. The same number of components  $n$  have vectors  $Y$ ,  $\beta$  and  $\epsilon$ . Where  $Y$  is a vector of dependence variable and  $\epsilon$  is a vector of errors. The  $\beta$  is a vector of unknown population parameter. Then the statistical model for linear regression between two variables looks like subsequent system of equations [[Gujarati2009](#)]:

$$\begin{bmatrix} Y_1 & Y_2 & \dots & Y_n \end{bmatrix}_{n \times 1} = \begin{bmatrix} 1 & X_1 & 1 & X_2 & \dots & \dots & \dots & \dots \\ \beta_1 & \beta_2 & \dots & \beta_3 \end{bmatrix}_{n \times 2} + \begin{bmatrix} \epsilon_1 & \epsilon_2 & \dots & \epsilon_n \end{bmatrix}_{n \times 1}$$

With simplified matrix notation, where the equation consists of a systematic component and a stochastic component:

$$Y = X\beta + \epsilon$$

syst. comp. + stoch. comp.

The object of a linear regression model is to estimate the parameter  $\beta^{\wedge}$ . The most commonly used projection technique for estimation of population parameter  $\beta^{\wedge}$  is to minimize residuals in squared form [[Rachev2007](#)]. The vector of residuals could be expressed:

$$e = Y - X \beta^{\wedge}$$

The objective function for optimization could be then assumed as following

[[Goldberger1964](#)]:

$$(e^T e)^2 \Rightarrow \min. \quad (7)$$

It is obvious that the following relation is valid:

$e^T e = (Y - X \beta^{\wedge})^T (Y - X \beta^{\wedge}) = Y^T Y - 2 \beta^{\wedge T} X^T Y + \beta^{\wedge T} X^T X \beta^{\wedge}$ . Certainly, it needs to derive  $\frac{\partial}{\partial \beta^{\wedge}}$  with respect to  $\beta^{\wedge}$  to find the minimum of given function:

$$\frac{\partial}{\partial \beta^{\wedge}} e^T e = -2 X^T Y + 2 X^T X \beta^{\wedge} = 0. \quad (8)$$

To make sure it is really a minimum the second order conditions must be solved:

$$\frac{\partial^2}{\partial \beta^{\wedge} \partial \beta^{\wedge}} e^T e = 2 X^T X.$$

As long as the full rank of X is true the matrix is positive definite. Therefore, assumption of finding a minimum was correct. Solving the equation [8](#) will generate *normal equations*

[[Verbeek2008](#)]:

$$(X^T X) \beta^{\wedge} = X^T Y \quad (9)$$

And after all the desired parameter  $\beta^{\wedge}$  will be:

$$\beta^{\wedge} = (X^T X)^{-1} (X^T Y).$$

## Properties of Ordinary Least Square estimators

The properties of estimators can be derived through normal equations [9](#).

If  $Y = X \beta^{\wedge} + e$  then:

$$(X^T X) \beta^{\wedge} = X^T (X \beta^{\wedge} + e).$$

Which provides following relation:

$$X^T e = 0.$$

## In accordance with the notation above the characteristics of OLS are:

- There is no correlation between observed X and residuals.
- The predicted Y are uncorrelated with residuals.
- The sum of residuals is equal zero.
- The sample mean of residuals is zero.
- The mean of predicted and observed Y is equal, then  $\overline{Y^{\wedge}} = \overline{Y}$ ; [[Wooldridge1995](#)].
- The regression hyperplane goes over the observed values means, then  $\overline{X} = \overline{Y}$ ; [[Wooldridge1995](#)].

More about OLS properties could be found in Greene:2003. Stated characteristics should always hold true. Even though the properties of residuals are certain, there are no any informations about inquired parameter  $\hat{\beta}$ . To make any further suppositions about the real population parameter  $\beta$ , it is required to make some assumptions.

## For a classical linear regression model the following assumption are essential:

- Linearity in Parameters,

*The assumption requires that the dependent variable  $Y$  is a linear combination of regressors and the stochastic terms  $\epsilon$ ; [Andersen1998].*

- The rank of matrix  $X$  is full,

*The matrix of explanatory variables must be full rank. There are two essential conditions for this assumption. The number of observation  $N$  is larger then number of explanatory variables ( $K$ ) or at least equal to  $N$ ,  $N \geq K$  [Wooldridge1995].*

- An exogenous Explanatory Variable,

*There is no ability of regressor to explain the error terms  $\epsilon$ ; [Berry1993].*

- The Error Terms have to be Independent and Identically Distributed,

*Given assumption requires the expected value of Errors to be zero and simultaneously the variance of Errors must be constant,  $\epsilon_i \sim iid(0, \sigma^2)$  [Berry1993]. Actually, it is the assumption of homoscedasticity and no autocorrelation.*

- The distribution of Error Terms in population is Normal.

*The claim of Central Limit Theorem could be applied for the Assumption 5, i.e. if  $N$  is enough large, then estimated coefficient will be asymptotically normal distributed [Gujarati2009].*

According to the Gauss-Markov Theorem, if the assumptions 1- 4 are satisfied, the Estimator of Ordinary Least Square is the Best Linear, Unbiased and Efficient Estimator (BLUE) [Greene2003].

In order to find the optimal hedge ration  $h^*$  a regression of spot on futures was provided. Since the closing prices do not fulfill the required assumption of classical linear regression the data was transformed. Instead of closing prices were applied the percentage changes of spot and futures. Thus, the optimal hedge ration according to OLS methodology should corresponds to estimated parameter  $\hat{\beta}$ . The transformed data are:

$$r_{st} = \ln(P_t \text{ spot} / P_{t-1} \text{ spot}), \quad (10)$$

and

$$r_{ft} = \ln(P_t \text{ futures} / P_{t-1} \text{ futures}). \quad (11)$$

P spot is the closing price of spot and P futures is the closing price of futures. Then the hedge ratio could be estimated by the following model:

$$r_{st} = \alpha + \beta r_{ft} + \epsilon_{st}$$

## 2.2 Naive portfolio

The hedging based on naive portfolio, or in other words equally-weighted portfolio, is the simplest way how one could protect against the price risk. The formal notation for weights of a naive portfolio could be expressed as:

$$w_s = w_f,$$

$$\text{or } w_f w_s = 1,$$

$$\text{hence } h^* = 1.$$

While it may seem that the naive portfolio is a primitive technique with insufficient hedging effect, so this may not be true. In favor of this approach speak several facts. The methodology is trivial and thus easily implementable. It is not too sensitive of small changes in the parameters comparing to more complex models. It is also more robust compared to other technique. In some cases the performance of naive portfolio is nearly so good as more sophisticated models [DeMiguel2009]. In different perspective, no model can always provide consistently better performance than a naive portfolio [Tu2011]. Some authors even suggest that a naive portfolio can achieve the highest Sharpe ration compared with other models [Poitras2002]. Moreover, arbitrarily chosen weights could serve as a benchmark [Amenc2002].

## 2.3 Error Correction Model

Financial time series are characterized by the presence of non-stationarity. The classical linear regression model can not be applied on such a data set. If a regression would be performed, then the results will be not correct. Although a model can show a significant statistical dependency, since the value of R<sup>2</sup> is high. In fact the model does not display a real dependency. The result shows rather on spurious regression [Greene2003]. The fundamental problem in financial time series is that the data is integrated, but not relate. Even [Granger1974] pointed out the problem with spurious regression.

Numerous analyzes have shown that economic and financial data can evince short-run and long-run relationships. A short-run relation exists only for a short time period and then disappears. In contrast, a long-run relation will not disappear over time. This relationship is described as an equilibrium and time series show a tendency to oscillate around it [Pesaran1996]. As the system is exposed to continual shocks it is never in equilibrium, but may be in the long equilibrium, i.e. in a state that converges over time to equilibrium. Such time series are then co-integrated [Engle1987].

The authors suggested two stepwise procedure to solve the problem with spurious regression. In the first step a stationary of researched data is investigated. The unit root test is applied for this purpose. The test of unit root constitutes to validate a hypothesis of random walk against an alternative hypothesis represented by the AR(1) process [Dickey1979]. Notation for unit root test according to given methodology is:

$$y_t = \theta y_{t-1} + \epsilon_{yt}$$

with the null hypothesis  $H_0 : \theta = 1$  and the alternative hypothesis  $H_1 : \theta < 1$ .

An alternative expression for the model:

$$\Delta y_t = (\theta - 1) y_{t-1} + \epsilon_t; \quad \epsilon_t = \delta y_{t-1} + \epsilon_t, \quad (12)$$

where  $\delta = \theta - 1$ . Thus the unit root hypothesis  $H_0 : \delta = 0$  against  $H_1 : \delta < 0$ .

However, the original Dickey-Fuller test carries certain disadvantages [Phillips1990]. Therefore a modified version so called Augmented Dickey-Fuller unit root test was introduced [Said1984]. Dickey-Fuller tests for time series with stationary and invertible residuals could be modified to:

$$y_t = \theta y_{t-1} + \epsilon_t,$$

where

$$\epsilon_t + \sum_{i=1}^p \varphi_i \epsilon_{t-i} = e_t + \sum_{j=1}^q \psi_j e_{t-j}, \quad (13)$$

and  $e_t \sim \text{IID}(0, \sigma^2)$ .

Although the lags of  $p$  and  $q$  values are unknown, the process could be approximated due an autoregressive process [Said1984]. Hence the unit root test could be based on the following model:

$$\Delta y_t = \theta y_{t-1} + \sum_{i=1}^n \psi_i \Delta y_{t-i} + \eta_t.$$

Afterwards testing the hypothesis  $H_0 : \theta = 0$ , which is the evidence of random walk in  $y_1, y_2, \dots, y_N$  and hence the process has a unit root. The alternative hypothesis expects  $H_1 : \theta < 1$ , thus stationary.

Once the order of differentiation to obtain a stationary in data is determined, it may be provided a regression to identified the long-term relationship in the series, or long memory -  $\beta$  [DeBoef2001]. The co-integrating regression corresponds thereto:

$$y_t = \alpha + \beta x_t + \mu_t. \quad (14)$$

The static cointegration regression is alike the OLS the closing prices of spot as dependent variable on futures closing prices as independent variable. Whenever a short-term relationship (short memory) within the residuals is demonstrated, that implies cointegration in time series [DeBoef2001]. Then the further changes in dependent variable  $\Delta y_t = y_t - y_{t-1}$  can be regressed on changes in independent variable  $\Delta x_t = x_t - x_{t-1}$  and the equilibrium error from the previous period. The model corresponds to:

$$\Delta y_t = \alpha + \beta \Delta x_t - \gamma \mu_{t-1} + \eta_t.$$

The parameter  $\beta$  represents an estimate for equilibrium rate and the parameter  $\gamma$  a short-run dynamics [DeBoef2008]. The model has to evince a permanent memory, in other words the unit root are confirmed and the errors from cointegrating regression 14 are not serially correlated and there will be no simultaneity [Enders1998]. The parameter  $\beta$  can be used as a hedge ratio  $h^* = \beta$  [Moosa2003].

## 2.4 Arch/Garch

Financial time series are characterized by the presence of a dynamic development of character traits. One of the most investigated field in finance is risk. Since, volatility represents a numerical measure of the risk, it attracted attention of scientists. Variation in the volatility of

financial data over time has been clearly demonstrated [[Andersen1997](#)]. In addition, it is possible to identify a volatility clustering in financial time series [[Lux2000](#)]. It was confirmed that the GARCH models could capture the volatility clustering [[Andersen1998](#)].

There are valid objections in the application of the standard OLS model for estimating the hedge ratio as the model is inconsistent. The inappropriateness emerges, because the model does not respect the heteroscedasticity embraced in prices [[Park1987](#)]. Another disadvantage of OLS model is a neglect of relevant information [[Myers1989](#)]. Given deficiencies were removed by introducing of innovative approaches in the form of autoregressive conditional heteroscedastic model (ARCH). The Arch model introduced by [[Engle1982](#)] should capture in the proper way the characteristic of heteroscedasticity in financial series, especially the variation of volatility and volatility clustering. The Arch model expresses the mean process of a financial time series in following way:

$$r_t = \mu + \epsilon_t,$$

where  $t=1,2, \dots, N$ ,  $r_t$  are the analyzed time series with  $N$  observations,  $\mu$  is the mean of the time series and  $\epsilon_t$  are residuals. The variances of residuals in ARCH model are assumed as:

$$\epsilon_t = \sigma_t z_t,$$

simultaneously  $z_t \sim N(0,1)$ , thus the  $\sigma_t^2$  corresponds to the following process:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_q \epsilon_{t-q}^2, \quad (15)$$

assuming  $\alpha_0 > 0$   $\alpha_i > 0$  and  $i > 1$ .

The ARCH model was able to describe the stochastic process of analyzed time series and to predict residuals. However, the model exhibits some shortcomings as well. Namely, the model works with a hypothesis of symmetrical shocks. In fact, a different effect on prices was confirmed by positive and by negative shocks [[Sadorsky1999](#)]. Further, the model could capture the volatility dynamics only if a sufficient number of observations and parameter are integrated [[Maddala1992](#)].

In order to exceed the restraints of ARCH model a generalized form of ARCH (GARCH) was introduced by Bollerslev. The conditional variance in the univariate GARCH model could be denoted as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_q \epsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_{t-p}^2, \quad (16)$$

on conditions  $\alpha_0 > 0$  and  $\beta_i > 0$ .

Unlike [15](#) the formula [16](#) numerous residual's lags can be replaced by limited number of conditional variances as well as the number of estimated parameters can be omitted [[Bollerslev1986](#)]. In fact, there is a large set of GARCH models [[Bollerslev2008](#)]. However, the most frequently used is obvious the basic form GARCH (1,1). The expression is based on the given formula [16](#):

$$\sigma_n^2 = \gamma V_L + \alpha \epsilon_{n-1}^2 + \beta \sigma_{n-1}^2,$$

where  $V_L$  represents a long-run variance. The parameters  $\gamma$ ,  $\alpha$ ,  $\beta$  are weights satisfying the condition:  $\gamma + \alpha + \beta = 1$ . In other words the  $\sigma_n^2$  is based on the recent term  $\epsilon_{n-1}^2$  and the recent  $\sigma_{n-1}^2$ . Substituting the  $\gamma V_L$  by  $\omega$  leads to mostly give form of GARCH (1,1):

$$\sigma_n^2 = \omega + \alpha \epsilon_{n-1}^2 + \beta \sigma_{n-1}^2. \quad (17)$$

As mentioned in the literature review the GARCH models were used for hedging purposes too. The main motivation for application of GARCH for hedging was argued by the same information set affecting both spot and futures prices [Baillie1991]. Hence, the bivariate GARCH (BGARCH) models have been used on cash and futures to estimate hedge ratio [Park1995]. The main benefit of the Multivariate GARCH application is its ability to capture time-varying hedge ratio.

The general formula of BGARCH for spot and futures could be expressed:

$$r_{st} = \mu_s + \epsilon_{st},$$

$$r_{ft} = \mu_f + \epsilon_{ft},$$

simultaneously:

$$[\epsilon_{st} \ \epsilon_{ft}] \sim N(0, H_t), \text{ and } H_t = \begin{bmatrix} h_{ss,t} & h_{sf,t} \\ h_{sf,t} & h_{ff,t} \end{bmatrix}.$$

Here  $r_{st}$  corresponds to [10](#),  $r_{ft}$  corresponds to [11](#),  $H_t$  is a positive definite matrix with conditional time-varying covariances.

Consequently the equation for  $H_t$  will be:

$$\text{vech}(H_t) = \text{vech}(C) + \sum_{i=1}^q \Gamma_i \text{vech}(\epsilon_{st}; \epsilon_{ft}) + \sum_{i=1}^p \Delta_i \text{vech}(H_{t-1}),$$

where the  $C$  is  $2 \times 2$  matrix and  $\Gamma_i$  and  $\Delta_i$  are  $3 \times 3$  matrices.

To ensure positive definiteness of the conditional covariance matrix is not feasible as long as non-linear restraints are not set up [Lamoureux1990]. In addition to that, the model assumes a large number of parameters. Which is providing a clumsy model. For this reason it is useful to make certain assumptions. Some simplification of the model will assume that  $H_t$  depends only on its lagged values and lagged residuals [Bera1997]. Thus, the model corresponds to:

The required condition for the model are  $C_{ss} > 0$ ,  $C_{ff} > 0$ ,  $C_{ss}C_{ff} - C_{sf}^2 > 0$ ,  $\gamma_{ss} > 0$ ,  $\gamma_{ff} > 0$ ,  $\gamma_{ss}\gamma_{ff} - \gamma_{sf}^2 > 0$ .

Another simplification of  $H$  is assuming that the conditional correlation among epsilon is constant [Bollerslev1990]. Such a model can be expressed as follows:

$$H_t = \begin{bmatrix} h_{ss,t} & h_{sf,t} \\ h_{sf,t} & h_{ff,t} \end{bmatrix} = \begin{bmatrix} h_{s,t} & 0 \\ 0 & h_{f,t} \end{bmatrix} \begin{bmatrix} 1 & \rho_{sf} \\ \rho_{sf} & 1 \end{bmatrix} \begin{bmatrix} h_{s,t} & 0 \\ 0 & h_{f,t} \end{bmatrix}. \quad (18)$$

The correlation coefficient thus does not depend on time in addition  $\rho_{sf} < 1$  alongside the parameters  $h_{s,t}$  and  $h_{f,t}$  are standard univariate GARCH process [17](#). The model modification is quite pleasant, however it is questionable whether the empirical data are consistent with the assumption of correlation [Bera1997].

Other adaptation of the model was introduced by [Engle1995]:

$$H_t = \begin{bmatrix} c_{ss} & c_{sf} & c_{ff} \end{bmatrix} + \begin{bmatrix} \gamma_{ss} & \gamma_{sf} & \gamma_{ff} \end{bmatrix} T \begin{bmatrix} \epsilon_{ss,t-1} & \epsilon_{sf,t-1} & \epsilon_{ff,t-1} \\ \epsilon_{sf,t-1} & \epsilon_{ss,t-1} & \epsilon_{ff,t-1} \\ \epsilon_{ff,t-1} & \epsilon_{sf,t-1} & \epsilon_{ss,t-1} \end{bmatrix} + \begin{bmatrix} \delta_{ss} & \delta_{sf} & \delta_{ff} \end{bmatrix} T H_{t-1} \begin{bmatrix} \delta_{ss} & \delta_{sf} & \delta_{ff} \end{bmatrix}.$$

The model guarantees a positive definiteness of the matrices. In the case that  $\Gamma$  and  $\Delta$  are 0, the  $H_t$  becomes a constant conditional covariance [Myers1989]:  $H_t = \begin{bmatrix} c_{ss} & c_{sf} & c_{ff} \end{bmatrix}$ .

For the estimation of the hedge ratio the presumption in [18](#) will be applied. Thus, the hedge ratio will be generated by BGARCH process and the expression states for:

$$h_t^* = h_{sf,t} / h_{ff,t}.$$



## 2.5 Wavelet

The scientific attention in the processing of financial time series was for a long time focused only on mathematical or statistical methods. However, in the last years some innovation in finance could be observed. One of new area in finance is application of tools for signal analyzing. The spectral analysis and Fourier transformation could be assigned as one of the unconventional tools applied in processing financial data [[Box2015](#)].

From the perspective of spectral analysis it is possible to examine the frequency behavior in the time series. A motivation why signal processing methodology may be applied could be the predictive ability. The results of the spectral analysis can be used in various areas. Among others, in risk management [[Acerbi:2002](#)]. The utilization was based on the findings of previous researches. In physics and astronomy the spectral analysis examines a light signal and its decomposition. It could be assumed that in the financial time series may occur similar structural changes corresponding to different frequencies [[Ozaktas2001](#)].

Usually modified data is used for the spectral analysis, i.e. data that is stationary. In principle this methodology composites the input data into a set of frequency bands [[Bloomfield2004](#)]. The interdependence of data is examined on the level of evaluated spectrum. The principal disadvantage of the approach is that it can not match the structural changes to the point where it was held [[Nason:1995](#)]. In other words, the examined data is decomposed without the time component.

<sup>10</sup>

However, for the examination of financial time series it represents a major shortage [[Gencay2001](#)].

Nevertheless, as mentioned the spectral analysis can work only with stationary data. If the data is not stationary, then it is necessary to transform it. This may cause loss of important information. Fortunately, there is a tool in the field of signal processing that can eliminate mentioned shortcomings. The tool is called wavelet transformation. The wavelets are functions with specific claims [[Vidakovic:2009](#)].

<sup>11</sup>

Similarly, like the Fourier Transform any function could be depicted via the sine and the cosine, any function could be represented by wavelets. Unlike the Fourier Transform the wavelet analysis maintains the time component. Since, the wavelet function has time-frequency localization property [[Chan1999](#)]. It can process non-stationary data and it is able to examine the interdependence in data [[Percival:2008](#)]. These assumptions make the wavelet appealing apparatus for application in hedging.

Analyzing the data via wavelet methodology especially in the field of finance is a relatively new discipline. However, the methodology is not a result of ultimate scientific contribution. According to the documented literature the concept of wavelet dates back to the beginning of the previous century [[Graps:1995](#)]. The theoretical foundations can be noted in the work of [[Haar1910](#)].

<sup>12</sup>

In the next decades the concept was not developed, nor was widely applied. The turn occurs in the 80's of the previous century due to the merit of Jean Morlet and other physicists and mathematicians.

<sup>13</sup>

There are distinct types of wavelet. Actually wavelets are functions which are applied to data or some other functions. The reason for using wavelets is that data after linear decomposition could be analyzed more efficiently. Although the wavelet transform guarantees a linear

expression of other function, the wavelets are non-linear functions [Ghanem:2001]. Thus, a function could be composited:  $f(t) = \sum_k \alpha_k \psi_k(t)$ .

Here  $k$  represents an integer for finite sum, eventually infinite sum,  $\alpha_k$  are the coefficients for expansion of real value and  $\psi_k$  is an expansion set representing a group of real-valued functions [Percival:2008].

In a speech of wavelet a signal or a function breaks down as follows:  $x(t) = \sum_k s_{J,k} \phi_{J,k}(t) + \sum_k d_{J,k} \psi_{J,k}(t) + \sum_k d_{J-1,k} \psi_{J-1,k}(t) + \dots + \sum_k d_{1,k} \psi_{1,k}(t)$ , (19)

where the function  $\phi(\cdot)$  is so called father wavelet and the function  $\psi(\cdot)$  stands for mother wavelet function, with the coefficients  $S_{J,k} = \int \psi_{J,k}(t) x(t) dt$  and  $d_{J,k} = \int \psi_{J,k}(t) x(t) dt$ ,  $j=1,2,3 \dots, J$  and concurrently the number of scales  $J = \log_2 n$ , where  $n$  is the number of data points and  $k$  ranges from 1 to the number of coefficients in given component. For more information see [Percival:2008]. So the concept of wavelet is based on multi-scale decomposition, multi-scale analysis or multi-resolution. The decomposition is intended into Hilbert space  $L^2(\mathbb{R})$  [Daubechies1992]. As indicated by Francis and Sangbae a two-dimensional family of functions is reflected in basic scaling function using scaling and translation on following manner [In2013]:  $\phi_{J,k}(t) = 2^{-j} \phi(2^{-j}t - k)$ . (20)

The expression  $2^j$  represents sequence of a scales, so is called scale factor and  $2^j k$  representing a translation or shift parameter. The scale factor is partitioning the frequency. The scaling function bridges a space vector over  $k$ :  $S_j = \text{Span}\{\phi_k(2^j t)\}$ .

Then the next conditions has to be met:  $\dots \subset S_{-2} \subset S_{-1} \subset S_0 \subset S_1 \subset \dots \subset L^2$ . (21)

Hence, the multi-resolution analysis allows to analyze time series into each of the approximation subspaces  $S_j$ . The multi-resolution equation represents the following relation:  $\phi(t) = \sum_k g(k) 2^{-j} \phi(2^{-j}t - k)$ ,  $k \in \mathbb{Z}$ , (22)

where  $g(k)$  is low-pass filter or scaling function coefficients and represents a sequence of real or complex numbers [Burrus1997]. Determining the scaling function is the first step to establish the wavelets, since their could be considered as a weighted sum of shifted scaling functions. Then the relation is declared to:  $\psi(t) = \sum_k h(k) 2^{-j} \phi(2^{-j}t - k)$ , (23)

where  $h(k)$  is a high-pass filter. Subsequently the mother function corresponds to:  $\psi_{j,k}(t) = 2^{-j} \psi(2^{-j}t - k) = 2^{-j} \psi(t - 2^j k)$ . (24) But then must be inevitably true that any time series  $x(t) \in L^2$  is possible to express as a series expansion in terms of scaling function and wavelets 19:  $f(t) = \sum_{k=-\infty}^{\infty} s(k) \phi_k(t) + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} d(j,k) \psi_{j,k}(t)$ .

And thus confirmed the property that any function can be expressed by a linear combination of wavelet or the scaling function respectively. However, for applying wavelet concept some conditions must be satisfied. According to [Mallat1989] the multi-resolution analysis could be applied on all square integrable functions in space  $L^2$ . Further, the assumptions include admissibility, vanishing moments and orthogonality [Kim:2005]. The first prerequisite is required more in theoretical applications.

14

The vanishing moments is satisfied from the 24. The orthogonality is fundamental for the wavelet transformation [Grossmann1984]. From the relation of the given father 20 and mother 24 functions the orthogonality is satisfied:  $\langle \phi(\cdot - k), \phi(\cdot - l) \rangle = 0, k, l \in \mathbb{Z}$ . For the continuous wavelet transformation the assumption of orthogonality would be expressed according to [Vidakovic2009]:  $\int \psi_{j,k} \cdot \psi_{\tilde{j},\tilde{k}} = 0$ , where  $j = \tilde{j}$  and  $k = \tilde{k}$  do not satisfied

contemporaneously. And finally having regard to the assumption in [21](#) the orthogonality must satisfy [[Tang2010](#)]:  $L^2 = S_0 \oplus D_1 \oplus D_2 \oplus D_3 \dots$

Apparently the affinity of  $S_0$  to the wavelet space corresponds to:  $S_0 = D_{-\infty} \oplus \dots \oplus D_{-1}$ .

Alternatively, the wavelet could be used for determining the low-pass filter from [22](#) and high-pass filter from [23](#). Thus, a continuous functions may break out into:  $g(h) = \int \varphi(t) \varphi(2t-k)$  &InvisibleTimes; &DifferentialD; t,  $h(k) = \int \psi(t) \psi(2t-k)$  &InvisibleTimes; &DifferentialD; t,

eventually:  $h(k) = (-1)^k g(k)$ .

Even with the usage of low-pass and high-pass filter it is essential to ensure the main assumptions, which are the value of mean:  $\sum_{k=0}^{J-1} h_k = 0$ ,

the prerequisite of unit energy:  $\sum_{k=0}^{J-1} h_k^2 = 1$ ,

and the already discussed orthogonality:  $\sum_{k=0}^{J-1} h_k h_{k+2n} = 0, n \in \mathbb{Z} \setminus \{0\}$ .

For the purpose of the thesis [1](#) will be applied to find the optimal hedge ratio. With respect to the sample size restriction, which must be a multiple integer of  $2^J$ , the maximum overlap discrete wavelet transform (MODWT) will be considered. For the hedging purpose it is necessary to find the variance of decomposition spot prices and covariance of decomposition spot and futures prices. If there is a stochastic process  $\{X\}$  and the sample size is divisible by  $2^J$  then applying the discrete wavelet transform the wavelet coefficient can be expressed from the high-pass filter in pyramid algorithm [[Mallat1989](#)]:

$$d_{j,t} = \sum_{k=0}^{J-1} h_{j,k} X_{t-1},$$

and the scaling coefficient analogous from the low-pass filter:  $s_{j,t} = \sum_{k=0}^{J-1} g_{j,k} X_{t-1}$ .

Apparently, there will be  $N/2^j$  scaling and wavelet coefficients. However, the compliance of discrete wavelet transform (DWT) conditions is rather complex [[Serroukh2000](#)]. Therefore the MODWT seems to be more convenient for released orthogonal assumption. Thus, the wavelet and scaling coefficients are [[Percival:2006](#)]:  $\tilde{d}_{j,t} = \frac{1}{\sqrt{2^j}} \sum_{k=0}^{J-1} \tilde{h}_{j,k} X_{t-1}$ ,

and  $\tilde{s}_{j,t} = \frac{1}{\sqrt{2^j}} \sum_{k=0}^{J-1} \tilde{g}_{j,k} X_{t-1}$ .

Thence the wavelet and the scaling filters are obtained from [[Percival:2006](#)]:  $\tilde{h}_{j,k} = h_{j,k} \sqrt{2^j}$ ,

and  $\tilde{g}_{j,k} = g_{j,k} \sqrt{2^j}$ .

Consequently, it is possible to write down the variance. The wavelet variance at scale  $j$  has following relation with the stochastic process  $\{X\}$ :  $\sum_{j=1}^{\infty} \sigma_{X,j}^2 = \sigma_X^2$ .

It is obvious that  $\sigma$  represents the contribution of variance at scale  $j$  to the overall variance. Given property enable to decomposition the variance into components that are associated to certain time scales [[Gallegati:2008](#)].

Hence, according to MODWT assumptions an unbiased estimator of the wavelet variance may be obtained:  $\tilde{\sigma}_{X,j}^2 = \frac{1}{N} \sum_{T=L_j}^N \tilde{d}_{j,t}^2$ ,

where  $N \sim = N - L_j + 1$  represents the maximal overlap coefficients at scale  $j$ . Further, the length of the wavelet filter at scale  $j$  is  $L_j = (2^j - 1)(L - 1) + 1$  [[Craigmile2005](#)]. To determine the futures/spot ratio the covariance is needed. The formula for wavelet covariance of two random variable  $X$  and  $Y$  at the scale  $j$  thusly can be expressed according to [[Vannucci1999](#)]:

$$\sigma_{XY} = \frac{1}{N} \sum_{t=L_j-1}^{N-1} \tilde{d}_{j,t} X \tilde{d}_{j,t} Y.$$

## 2.6 Copula

One of the most important financial topic or issue in risk management is the quantification of overall risk. In other words the aggregation of individual risks. However, a simple summation should be considered only if the assets are not dependent. Likewise, the relation must not be strictly linear. In such a case the expression of the joint risk become more complex, because the joint distribution is mostly unknown. If there should be expressed a wider set of dependencies between multiple assets commonly a Gaussian behavior is assumed. Then the information of dependences is displayed in a covariance matrix.

The investigation of dependence in quantitative interpretation goes deep into history. According to the preserved information already in the seventeenth century data maps were applied to describe monsoon rains [Dorey2005]. However for the field of finance the introduction of the correlation concept in 1888 by Francis Galton was crucial [Hauke2013]. After [Pearson1896] worked out the concept of bi-variate normal correlation in 1896.

16

The Pearson moment-product correlation corresponds to:  $\rho_{XY} = E[(X - \mu_X)(Y - \mu_Y)] / \sigma_X \sigma_Y$ . (25)

Since the Pearson product moment correlation can be used only if some strong requirements are met, an another methodology that relaxes these assumptions was introduced, so called Spearman rank correlation [Spearman1904]. Thus, 25 could be applied on ranked data or expressed by own formula:  $r_s = 1 - \frac{6 \sum_{i=1}^n (p_i - q_i)^2}{n(n^2 - 1)}$ , where  $p_i$  and  $q_i$  are ordered value from any random variables. The Spearman's model is a non-parametric measurements of dependence. It reflects rather the strength of monotonic relations. A similar concept based on ordered data represents the rank correlation introduced by [Kendall1948]. The formula then corresponds to:  $\tau = \frac{C_p - D_p}{n(n-1)/2}$ ,

here  $C_p$  stands for concordant pairs and  $D_p$  represents the number of discordant pairs from the random variable  $X$  and  $Y$ , with  $n$  pairs  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ . Thus, a concordant pair is present if  $x_i > y_j \wedge y_i > y_j$  or a pair is said to be discordant if  $x_i > y_i \wedge x_i < y_j$  or  $x_i < y_i \wedge x_i > y_j$  and simultaneously  $i \neq j$ .

Nevertheless, until the fifties almost the Pearson correlation was applied in finance. Moreover, the relation was examined as a cross-section data analysis. However, the problem of temporal correlation began to be gradually more significant [Longin1995].

Many of financial and econometric models are based on strong assumptions about the distribution. Usually it is assumed a normal distribution. On these assumptions are based as well the main financial tools like the Capital Asset Pricing Model and the Arbitrage Pricing Theory. However, the real data mostly does not fulfill such requirement. For instance, the asset returns are fat-tailed [Rachev:2005]. Additionally, if dependence of two random variables is examined, it is assumed that they have identical distributions. But empirical observations disprove such assumption often.

The above mentioned shortcomings were initiating new developments of risk models. Among others, the copula concept, which was introduced to the field of finance in the beginning of the 21. century. The aim of scientists was to highlight the deficiencies in the application of Pearson's correlation [Bouye2000]. Apart from normal distribution it included the inability to work with time-varying volatility. Longin1995 Eventually, it is not possible to deal correctly with the problem of heteroskedasticity in data. Loretan2000 Furthermore, some articles

pointed out the relationship in extreme values [[Embrechts2001](#)]. Primarily, the ability of copula to describe the relation between assets in extreme events like the period of crises advocated the application of copula [[Rockinger2001](#)], [[Hartmann2004](#)].

17

Especially, the crises caused by balance of payment [[Costinot2000](#)]. The copula gained a considerable popularity in the application of contagion [[Rodriguez2007](#)]. Paradoxical one of the model from copula family has been blamed as one of the causes of the recent global financial crisis [[Salmon2012](#)].

18

The birth of the statistical tool copula can be found in the fifties of the last century. The basic of copula were given presumably in contrition of the work on bivariate and trivariate distributions with given univariate margins [[Frechet1951](#)], [[Dall1956](#)]. Nevertheless, the real breakthrough was the work of [[Sklar1959](#)]. He introduced a new function with the name copula, that joins multiple distribution functions to joint one-dimensional marginal distribution function [[Nelsen1991](#)].

19

The initial impetus for finding copula was study on probabilistic metric space on theoretical level [[Sklar1959](#)]. The first statistical application realized first in eighties [[Schweizer1981](#)]. Some scientific articles in this period dealt with the question of whether there exist a linkage between above given metrics of dependence [[Frees1998](#)].

20

In fact there are more copula functions. Actually, copulas represent instruments which can describe the dependence properties of a multivariate random vectors and afford to identify the relation between the joint distribution and the marginal distributions. Consider a pair of continuous random variables ( X,Y ) with their marginal cumulative distribution functions:

$$F(x)=P( X \leq x ) \text{ and } G(y)=P( Y \leq y ),$$

such that  $\forall x,y \in \mathbb{R}$  and a joint cumulative distribution function:

$$H( x,y )=P( X \leq x, Y \leq y ),$$

here  $F(x)$  ,  $G(y)$  are marginal cumulative distribution functions both in interval  $I = [ 0,1 ]$  and  $H( x,y )$  is a joint cumulative distribution function in  $I$  . According to Sklar Theorem if there is a joint distribution function  $H( x,y )$  , then there exists a copula  $C(u,v)$  with the bivariate uniform distribution [[Sklar1959](#)]:  $C( u,v )=P( U \leq u, V \leq v ), u,v \in [ 0,1 ]$  .

Furthermore, the joint distribution function could be expressed through the Copula:  $H( x,y )=C( F(x),G(y) )=C( u,v )$  .

Thus, there are transformed variables  $U=F(X)$  and  $V=G(Y)$  both in the interval  $I$  .

Conversely, the copula could be defined by the inversion distribution function:  $C( u,v )=H( F^{-1}( u), G^{-1}( v) )$ , (26)

where  $F^{-1}$  is the pseudo-inverse of  $F$  and  $G^{-1}$  is the pseudo-inverse of  $G$  or the quantile functions of margins. In this point of view copula transforms the random variables  $X$  and  $Y$  into other random variable  $( U,V )=( F(x),G(y) )$  , which have the margins in  $I$  . But it is important that the dependence will be preserved. In other words, copula allows to analyzed the dependence within random variables with distinct marginal distributions.

It is apparent that the copula function is from the domain  $I^2$  to  $I$  .

21

The copula function has to fulfill the consequent properties [[Nelsen2013](#)]:

- For every  $u, v$  in  $I$

22

$$C(u, 0) = 0 \text{ \&xwedge; } C(v, 0) = 0,$$

and  $C(u, 1) = u \text{ \&xwedge; } C(v, 1) = v$ .

- For every  $u_1, u_2, v_1, v_2$  in  $I$  such that  $u_1 < u_2$  and  $v_1 < v_2$ ,  $C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$ . This property guaranties that the  $C(\cdot)$  is 2-increasing or quasi-monotone.

23

- Copula is a Lipschitz function:  $|C(u_1, v_1) - C(u_2, v_2)| \leq |u_1 - v_1| + |u_2 - v_2|$ . This condition ensures that the copula function is continuous on its domain.

24

For the purposes of the thesis the random variable were closing spot and futures prices. Since, there are large number of copula functions it is cardinal first to select the appropriate copula function. The R-package VineCopula was applied to determined the corresponding copula on the three analyzed commodities [Brechmann2013]. The logarithmic returns were examined by all three commodities. Hence, the results selected the t-copula as the appropriate from the copula family in all three cases.

25

## Bivariate Student's t-copula

It would be advisable to begin with the canonical univariate Student's t distribution. The t distribution seems to be more appropriate to the real data then the normal distribution. The reason for this statement is, because the t distribution unlike the normal distribution has heavier tails [Ruppert2004]. Then the Student's t distribution is given by the probability density function  $f_{v,t}(x) : f_{v,t}(x) = \frac{\Gamma(v+1/2)}{\Gamma(v/2)} \pi^{-v/2} (1+x^2/v)^{-(v+1/2)}$ ,  $-\infty < x < \infty$ ,

with the parameter  $v$

26

identifying the degrees of freedom simultaneously it must hold  $v > 0$  and the  $\Gamma(\cdot)$  is the Euler Gamma function [Cherubini2011].

27

Thusly the bivariate correlated t distribution corresponds to:  $f_{2,v,t}(x,y) = \frac{1}{2} \pi^{1-\rho} (1+x^2+y^2-2\rho xy)^{-v} (1+v)^{-1}$ ,

here  $\rho$  is the correlation coefficient 25 [Winer1971]. Respecting 26 then the bivariate Student's t copula corresponds to the following equation:  $C_{\rho,v,t}(u,v) = \int_{-\infty}^{t_1} f_{v,t}(u) f_{v,t}(v) dt_1$

where  $f(t_i)$  represents the density function of Student's t distribution and  $t_{v-1}$  denotes the quantile function of the standard univariate Student's t distribution [Demarta2005]. Hence, the copula density function according to [Embrechts2001] is:  $c_{\rho,v,t}(u,v) = f_{\rho,v,t}(F_{v-1}(u), F_{v-1}(v)) f_{v,t}(F_{v-1}(u)) f_{v,t}(F_{v-1}(v))$ ,  $u, v \in I$ .

In the above equation  $F^{-1}(\cdot)$  denotes the quantile function of the marginal  $t$  distribution with  $\nu$  degrees of freedom and  $f_{\nu, t}$  is the joint density function.

The properties of  $T$  copula are determined by the underlying Student's  $t$  distribution:

- The copula is symmetric,
- It belongs to the elliptical copula family,

28

- The  $T$  copula exhibits tail dependence.

The hedge ratio is calculated from copula covariance matrix generated from simulated data. After the copula is selected, the required procedure for obtaining the covariance matrix could be initiated. The algorithm consist of sampling the multivariate Student's distribution with an appropriate correlation matrix  $R$  [[Fantazzini2004](#)]. Then each margin is converted employing the probability integral transformed with the  $t$  distribution function. According to [[Embrechts2001](#)] the algorithm supposes following steps:

- Find Cholesky decomposition  $L$  for the correlation matrix  $R$ ,

29

- Generate a vector  $Z$  with  $p$  independent random variables  $Z = (z_1, z_2, \dots, z_p) \in N(0, 1)$ ,
- Simulate a random variate  $s$  from  $\chi^2$  such that is independent of  $Z$ ,
- To obtain a  $p$ -variate normal random variable with correlation matrix  $R$ , it must be set  $y = LZ$ ,
- Set  $x = \sqrt{s} y$ ,
- Set  $u_i = t_{\nu}(x_i), i=1, 2, \dots, p$ , here  $t_{\nu}$  represents the univariate cumulative  $t$  distribution with  $\nu$  degrees of freedom,
- Then a sample of  $T$  copula with  $\nu$  distribution function and correlation structure  $R$  could be denoted as  $(u_1, u_2, \dots, u_p) \sim C(R, \nu, t)$ .

In the analysis of the thesis the number of iterations corresponded to the number of observations in examined data. Afterwards, the covariance matrix from the simulated data was used to calculate the optimal hedge ration in accordance with [1](#).

## 2.7 Mean Extended Gini Coefficient

The concept of mean-variance portfolio theory is based on the normality in returns [[Markowitz2000](#)].

30

However, the empirical data does not fulfill this assumption. Taking into account the absence of the attitude to risk aversion of decision makers, the stochastic-dominance rules can not be hold. Providing the required conditions in accordance with the stochastic-dominance theories is arduous, even when the assumptions of maximizing expected utility are fulfilled [[Chen2013a](#)]. The Gini's mean approach solves some shortages, which other models can not treat, and ultimately provides a framework respecting the stochastic-dominance theory [[Yitzhaki1983](#)]. The Gini's coefficient was initially used for the analysis of wealth

distribution in society [Gini1921]. However, this mathematical apparatus was latter applied on risk evaluation too [Yitzhaki1982].

Since, the Gini's methodology incorporates area for measuring a variability of random variable. From the perspective of portfolio theory the Gini's mean difference could be applied on random return. Thus, let  $R$  be a random return of a portfolio falling into the interval  $a, b$  range,  $F(\cdot)$  and  $f(\cdot)$  is the distribution function and the density function of  $R$ , respectively. Further,  $F(a)=0$  and  $F(b)=1$ . Then, if  $\Gamma$  is the sign for Gini's mean difference, the equation corresponds to:  $\Gamma = 1/2 E\{ |R_1 - R_2| \}$ ,

$$\text{alternatively: } \Gamma = 1/2 \int_a^b \int_a^b |r_1 - r_2| f(r_1)f(r_2) dr_1 dr_2 \quad (27)$$

Under a condition that both arguments  $R_1$  and  $R_2$  are independent and evince the same distribution like  $R$  [Lien1993]. Essentially,  $\Gamma$  can be estimated on following way:  $\Gamma = \int_a^b [1-F(r)]^2 dr - \int_a^b [1-F(r)]^2 dr$ .

Afterwards, the variance of  $R$  is:  $\Gamma = 1/2 E[(R_1 - R_2)^2]$ ,

$$\text{alternatively: } \Gamma = 1/2 \int_a^b \int_a^b (r_1 - r_2)^2 f(r_1)f(r_2) dr_1 dr_2$$

[Shalit1995]refined the original model and enhanced it by the measure of risk aversion. The new defined model, Mean-extended Gini (MEG), with the parameter of risk aversion  $v$ , where  $1 \leq v < \infty$  has following form:  $\Gamma(v) = \int_a^b [1-F(r)]^v dr - \int_a^b [1-F(r)]^v dr$ ,

and after modification according to [Lerman1984]:  $\Gamma(v) = \mu - a - \int_a^b [1-F(r)]^v dr$ .

In the case of a risk-neutral investor the parameter  $v=1$  and then  $\Gamma(1)=0$ . Equally, the situation  $v=2$  represents a special case of mean-extended Gini, where it corresponds to the Gini's mean difference [27], so  $\Gamma(2) = \Gamma$ . Obviously, if  $v$  increase indefinitely, the term will reduce to  $\lim_{v \rightarrow \infty} \Gamma(v) = \mu - a$ . Presumably, the process of computing  $\Gamma$  from the equation [27] will be onerous. Therefore, a viable solution for  $\Gamma(v)$  was proposed by [Shalit1984] in the following form:  $\Gamma(v) = -v \text{Cov}(R, [1-F(R)]^{v-1})$ .

The modification above allows a convenient calculation. The application of Extended mean Gini's Coefficient allows to satisfy the first and second degree of stochastic dominance [Hey1980]. Hence, it is a proper tool for hedging purposes. A proof of satisfying the conditions can be introduced by following statement.  $\lambda_n = \int_a^b [1-F(r)]^n dr - \int_a^b [1-G(r)]^n dr, n=1,2,3, \dots$ , (28)

where  $F(\cdot)$  and  $G(\cdot)$  are distribution functions of return in a portfolio A and a portfolio B. If B is dominated by A, then the necessary condition for the first and the second stochastic dominance are  $\lambda_n > 0, n=1,2,3, \dots$  [Yitzhaki1982]. After integration by part [28] and  $n=1$  the equation is:  $\lambda_1 = \mu_A - \mu_B$ ,

here  $\mu_A$  is the mean of portfolio A and  $\mu_B$  is the mean of portfolio B. Simultaneously, let  $\Gamma_A(v)$  and  $\Gamma_B(v)$  are the extended Gini's coefficients, so after that for  $n=2$  the expression get following form:  $\lambda_2 = (\mu_A - \Gamma_A(2)) - (\mu_B - \Gamma_B(2))$ . (29)

Then  $\mu_A > \mu_B$  guarantees the first-degree stochastic dominance and  $\mu_A - \Gamma_A(2) > \mu_B - \Gamma_B(2)$  guarantees the second-degree stochastic dominance [Levy1992]. Unless, the  $\Gamma(v)$  is used



as measurement of risk, then the optimal hedge ratio could be found by minimizing the  $\Gamma(V)$  [Kolb1992]. Although as evidenced by Kolb and Okunev the optimal hedge ratio can also be obtained by maximizing [Kolb1993].

31

[Shalit1995] examined the value of a portfolio. A rational investor prefers a larger value of portfolio before smaller. The value of a portfolio  $V_{pt}$  at a given time  $t$  could be expressed:  $V_{pt} = P_{st} + h(P_{ft-1} - P_{ft})$ ,

where  $P_s$ ,  $P_f$  are prices of spot and futures, respectively. [Shalit1995] determine the optimal hedge ratio by the following formula:  $h^* = \frac{\text{Cov}(P_s, [1-G(V)]^{v-1})}{\text{Cov}(P_f, [1-G(V)]^{v-1})}$ ,

taking  $G(\cdot)$  as a distribution function of  $V$  and  $r_s$  and  $r_f$  are returns of spot and futures.

Assuming that the distribution function of  $r_f$  is similar to  $G(V)$ , since the empirical ranking of  $V_p$  should be alike  $P_f$ , then the reliable estimate of  $h^*$  could be stated in the act of:  $h^* = \frac{\sum_{i=1}^n (r_{si} - \bar{r}_s)(z_i - \bar{z})}{\sum_{i=1}^n (r_{fi} - \bar{r}_f)(z_i - \bar{z})}$ ,

where  $z_i = [1-G(r_{fi})]^{v-1}$  [Lien2000]. Since  $v=2$  satisfies the conditions 29 so the calculation in the practical part based on MEG will consider only  $v=2$ .