VALUE AT RISK

Irina Khindanova
University of California, Santa Barbara
Economics Department

Latest version: April 27, 1998
Part A. Value at Risk (VAR): Importance, Existing Methodologies, and a Critique

1. Introduction: VAR and the New Bank Capital Requirements for Market Risk

One of the most important tasks of financial institutions is evaluation of exposure to market risks, which arise from variations in prices of equities, commodities, exchange rates, and interest rates. The dependence on market risks can be measured by changes in the portfolio value, or profits and losses. A commonly used methodology for estimation of market risks is the Value at Risk (VAR).

Regulators and the financial industry advisory committees recommend VAR as a way of risk measuring. In July 1993, the Group of Thirty first advocated the VAR approaches in the study “Derivatives: Practices and Principles”\(^1\). In 1993, the European Union instructed setting capital reserves to balance market risks in the Capital Adequacy Directive “EEC 6-93”, effective from January 1996\(^2\). It was an improvement with respect to the 1988 Basle Capital Adequacy Accord of G10, which centered on credit risks and did not consider market risks in details\(^3\).

In 1994, the Bank for International Settlements in the Fisher report advised disclosure of VAR numbers\(^4\). In the April 1995 proposal “Supervisory Treatment of Market Risks”, the Basle Committee on Banking Supervision suggested that banks can use their internal models of VAR estimations as the basis for calculation of capital requirements\(^5\). In January 1996, the Basle Committee amended the 1988 Basle Capital Accord\(^6\). The supplement suggested two approaches to calculate capital reserves for market risks: “standardized” and “internal models”\(^7\). According to the in-house models approach, capital requirements are computed from multiplying the banks’ VAR values by a factor between three and four. In August 1996, the US bank regulators endorsed the Basle Committee amendment\(^8\). The Federal Reserve Bank allowed the two-year period for its

---

1 Kupiec, 1995; Simons, 1996; Fallon, 1996; Liu, 1996
2 Liu, 1996
3 Jackson, Maude, and Perraudin, 1995; 1997
4 Hendricks, 1996
5 Kupiec, 1995; Jorion, 1996; Beder, 1995
6 Basle Committee on Banking Supervision, 1996
7 Simons, 1996; Jackson, Maude, and Perraudin, 1997; Hopper, 1997
8 Lopez, 1996
implementation. The proposal is effective from January 1998\(^9\). The US Securities and Exchange Commission suggested to apply VAR for enhancing transparency in derivatives activity. Derivatives Policy Group has also recommended VAR techniques for quantifying market risks\(^{10}\).

The use of VAR models is rapidly expanding. Financial institutions with significant trading and investment volumes employ the VAR methodology in their risk management operations\(^{11}\). In October 1994, JP Morgan unveiled its VAR estimation system\(^{12}\), RiskMetrics™. Credit Swiss First of Boston developed proprietary Primerisk and PrimeClear (March 1997). Chase Manhattan’s product is called Charisma. Bankers Trust introduced the RAROC in June 1996. Deutsche Bank uses the dbAnalyst 2.0 from January 1995. Corporations use VAR numbers for risk reporting to management, shareholders, and investors since VAR measures allow to aggregate exposures to market risks into one number in money terms. It is possible to calculate VAR for different market segments and to identify the most risky positions. The VAR estimations can complement allocation of capital resources, setting position limits, and performance evaluation\(^{13}\).

In many banks the evaluation and compensation of traders is derived from returns per unit VAR. Nonfinancial corporations employ the VAR technique to unveil their exposure to financial risks, to estimate riskiness of their cashflows, and to undertake hedging decisions. Primers of applying VAR analysis for estimating market risks by nonfinancial firms are two German conglomerates Veba and Siemens\(^{14}\). The Norwegian oil company Statoil implemented a system, which incorporates the VAR methodologies\(^{15}\). Corporations hedge positions to “buy insurance” against market risks. An appealing implication of VAR is as an instrument for corporate self-insurance\(^{16}\). VAR can be explained as the amount of uninsured loss that a corporation accepts. If the self-insurance losses are greater than the cost of insuring by hedging, the corporation should buy external insurance. Investment analysts employ VAR techniques in project

\(^{9}\) Hendricks and Hirtle, 1997
\(^{10}\) Kupiec, 1995
\(^{11}\) Heron and Irving, 1997; The Economist, 1997
\(^{12}\) JP Morgan, 1995
\(^{13}\) Liu, 1996; Jorion, 1996
\(^{14}\) Priest, 1997a and 1997b
\(^{15}\) Hiemstra, 1997
\(^{16}\) Shimko, 1997a
valuations\textsuperscript{17}. Institutional investors, for instance, pension funds, use VAR for quantifying market risks.

The new market risk capital requirements become effective from January 1998. The US capital standards for market risks are imperative for banks with trading accounts (assets and liabilities) greater than $1 billion or 10 percent of total assets\textsuperscript{18}. Though, the regulators can apply these standards to banks with smaller trading accounts. The market risk capital requirements allow to calculate capital reserves based either on "standardized" or "internal models" methods. The standardized method computes capital charges separately for each market (country) assigning percentage provisions for different exposures to equities, interest rate and currency risks\textsuperscript{19}. The total capital charge equals the sum of the market capital requirements. The main drawback of the "standardized" approach is that it does not take into consideration global diversification effects\textsuperscript{20}. The second approach determines capital reserves based on in-house VAR models. The VAR values should be computed with a 10-day time horizon at a 99 percent confidence level using at least one year data\textsuperscript{21}.

The new capital requirements classify market risk on general market risk and specific risk. The \textit{general risk} is the risk from changes in the overall level of equity and commodity prices, exchange rates and interest rates. \textit{Specific risk} is the risk from changes in prices of a security because of reasons associated with the security’s issuer.

The capital requirement for general market risk is equal to the maximum of:

(i) the \textit{current VAR} ($\text{VAR}_t$) number and

(ii) the \textit{average VAR} \( \left( \frac{1}{60} \sum_{i=1}^{60} \text{VAR}_{t-i} \right) \) over the previous 60 days multiplied by a factor between three and four.

The capital charges for specific risk cover debt and equity positions. The specific risk estimates obtained from the VAR models should be multiplied by a factor of four. Thus, a market risk capital requirement at time $t$, $C_t$, is

\textsuperscript{17} Shimko, 1997b
\textsuperscript{18} Hendricks and Hirtle, 1997
\textsuperscript{19} Example: The required capital reserves for positions in the US market recognize hedging by the US instruments but do not consider hedging by the UK instruments.
\textsuperscript{20} In other words, the "standardized" method ignores correlations across markets in different countries. See Jackson, Maude and Perraudin, 1997; Liu, 1996
where $A_t$ is a multiplication factor between three and four, $S_t$ is the capital charge for specific risk.

The $A_t$ values depend on accuracy of the VAR models in previous periods. Denote by $K$ the number of times when daily actual losses exceeded the predicted VAR values over the last year, or the last 250 trading days. Regulators split the range of values of $K$ into three zones: the green zone ($K \leq 4$), the yellow zone ($5 \leq K \leq 9$), and the red zone ($K \geq 10$). If $K$ is within the green zone, then $A_t = 3$. If $K$ is within the yellow zone, $3 < A_t < 4$, in the red zone, $A_t = 4$.

A VAR measure is the highest possible loss over a certain period of time at a given confidence level. Example: The daily VAR for a given portfolio of assets is reported to be $2$ million at the 95 percent confidence level. This value of VAR means that, without abrupt changes in the market conditions, one-day losses will exceed $2$ million 5 percent of the time.

Formally, a $\tau=\tau_{VAR}$ is defined as the upper limit of the one-sided confidence interval:

$$\Pr[\Delta P(\tau) < -VAR] = 1 - \alpha$$

(1)

where $\alpha$ is the confidence level and $\Delta P(\tau) = \Delta P(\tau)$ is the relative change (return) in the portfolio value over the time horizon $\tau$.

\[\Delta P(\tau) = P(t+\tau) - P(t),\]

---

21 For the exact definition of VAR see (1) with $\tau = 10$ and $\alpha = .99$ later in this section.

22 The regulators recommend to use the time horizon $\tau$ of 10 days (two weeks) in VAR estimations. For backtesting, the regulators use $\tau = 1$ day.

23 For more detailed explanation of the time horizon and the window length see also 3.3 and 3.4.

24 Denote by $\hat{\kappa}$ the fraction of days, when the observed losses exceeded the VAR estimate. If $K=10$, then $\hat{\kappa}$ is $10/250 = 0.04$. However, the 99% confidence level implies probability of 0.01 for exceeding the VAR estimate of daily losses.
where $P(t+\tau) = \log S(t+\tau)$ is the log-spot value at $t+\tau$, $P(t) = \log S(t)$, $S(t)$ is the portfolio value at $t$, the time period is $[t, T]$, with $T - t = \tau$, and $t$ is the current time.

The time horizon, or the holding period, should be determined from the liquidity assets and the trading activity. The confidence level should be chosen to provide a rarely exceeded VAR value.

The VAR measurements are widely used by financial entities, regulators, non-financial corporations, and institutional investors. Clearly, VAR is of importance for practitioners and academia alike. The aim of this paper is to review the recent approaches to VAR and to outline directions for new empirical and theoretical studies.

In Section 2 we discuss traditional approaches to approximations of the distribution of $\Delta P$ and VAR computations. Section 3 analyzes components of VAR methodologies. The ways of verifying accuracy of methods and performance measures are analyzed in Section 4. Section 5 reports VAR strengths and weaknesses. Section 6 presents suggested VAR modifications and improvements. Section 7 outlines future research. Section 8 states conclusions.

### 2. Computation of VAR

From the definition of $VAR = VAR_{t,\tau}$, (1), the VAR values are obtained from the probability distribution of portfolio value returns:

$$1 - \alpha = F_{\Delta P}(-VAR) = \int_{-\infty}^{-VAR} f_{\Delta P}(x) dx,$$

where $F_{\Delta P}(x) = \Pr(\Delta P \leq x)$ is the cumulative distribution function (cdf) of portfolio returns in one period, and $f_{\Delta P}(x)$ is the probability density function (pdf) of $\Delta P$. The VAR methodologies mainly differ in ways of constructing $f_{\Delta P}(x)$.

The traditional techniques of approximating the distribution of $\Delta P$ are:

- the *parametric method* (analytic or models-based),
- *historical simulation* (nonparametric or empirical-based),
• Monte Carlo simulation (stochastic simulation), and
• the stress-testing (scenario analysis).\(^{26}\)

2.1. Parametric Method

If the changes in the portfolio value are characterized by a parametric distribution, VAR can be found as a function of distribution parameters. In this section we review:

• applications of two parametric distributions: normal and gamma,
• linear and quadratic approximations to price movements.

2.1.1. VAR for a Single Asset

Assume that a portfolio consists of a single asset, which depends only on one risk factor. Traditionally, in this setting, the distribution of asset return is assumed to be the univariate normal distribution, identified by two parameters: the mean, \(\mu\), and the standard deviation, \(\sigma\). The problem of calculating VAR is then reduced to finding the \((1-\alpha)\)th percentile of the standard normal distribution \(z_{1-\alpha}\):

\[
1 - \alpha = \int_{-\infty}^{X^*} g(x) dx = \int_{-\infty}^{z_{1-\alpha}} \phi(z) dz = N(z_{1-\alpha}), \text{ with } X^* = z_{1-\alpha}\sigma + \mu,
\]

where \(\phi(z)\) is the standard normal density function, \(N(z)\) is the cumulative normal distribution function, \(X\) is the portfolio return, \(g(x)\) is the normal distribution function for returns with the mean \(\mu\) and the standard deviation \(\sigma\), and \(X^*\) is the lowest return at a given confidence level \(\alpha\).

---

\(^{25}\) If \(f_{\Delta P}(x)\) does not exist, then VAR can be obtained from cdf \(F_{\Delta P}\).

Investors in many applications assume that the expected return $\mu$ equals 0. This assumption is based on the conjecture that the magnitude of $\mu$ is substantially smaller than the magnitude of the standard deviation $\sigma$ and, therefore, can be ignored. Then, it can be assumed:

$$X^* = z_{1-\alpha} \sigma.$$ 

and, therefore,

$$\text{VAR} = -Y_0 X^* = -Y_0 z_{1-\alpha} \sigma,$$

where $Y_0$ is the initial portfolio value.

2.1.2. Portfolio VAR

If a portfolio consists of many assets, the computation of VAR is performed in several steps. Portfolio assets are decomposed into “building blocks”, which depend on a finite number of risk factors. Exposures of the portfolio securities are combined into risk categories. The total portfolio risk is constructed, based on aggregated risk factors and their correlations. We denote:

- $X_p$ is the portfolio return in one period,
- $N$ is the number of assets in the portfolio,
- $X_i$ is the $i$-th asset return in one period $(\tau = 1)$, $X_i = \Delta P(1) = P_i(1) - P_i(0)$, where $P_i$ is the log-spot price of asset $i$, $i=1,...,N$. More generally, $X_i$ can be the risk factor that enters linearly\(^{27}\) in the portfolio return.
- $w_i$ is the $i$-th asset's weight in the portfolio, $i=1,...,N$.

$$X_p = \sum_{i=1}^{N} w_i X_i$$

In matrix notation,

\(^{27}\) If the risk does not enter linearly (as in a case of an option), then a linear approximation is used.
\[ X_p = w^T X, \]
where \( w = (w_1, w_2, \ldots, w_N)^T \),
\[ X = (X_1, X_2, \ldots, X_N)^T. \]

Then the portfolio variance is
\[
V(X_p) = w^T \Sigma w = \sum_{i=1}^{N} w_i^2 \sigma_{ii} + \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} w_i w_j \rho_{ij} \sigma_{i} \sigma_{j}
\]
where \( \sigma_{ii} \) is the variance of returns on the i-th asset, \( \sigma_{i} \) is the standard deviation of returns on the i-th asset, \( \rho_{ij} \) is the correlation between the returns on the i-th and the j-th assets, \( \Sigma \) is the covariance matrix, \( \Sigma = [\sigma_{ij}], 1 \leq i \leq N, 1 \leq j \leq N. \)

If all portfolio returns are \textit{jointly normally distributed}, the portfolio return, as a linear combination of normal variables, is also \textit{normally distributed}. The portfolio VAR based on the normal distribution assumption is
\[
\text{VAR} = -Y_0 z t - \alpha \sigma(X_p),
\]
where \( \sigma(X_p) \) is the portfolio standard deviation (the \textit{portfolio volatility}),
\[
\sigma(X_p) = \sqrt{V(X_p)}.
\]

Thus, risk can be represented by a combination of linear exposures to normally distributed factors.

In this class of parametric models, to estimate risk, it is sufficient to evaluate the covariance matrix of correlations of portfolio risk factors (in the simplest case, individual asset returns).

The estimation of the covariance matrix is based on the \textit{historical data} or on \textit{implied data} from securities pricing models.

If portfolios contain zero-coupon bonds, stocks, commodities, and currencies, VAR can be computed from correlations of these basic risk factors and the asset weights. If portfolios include more complex securities, then the securities are decomposed into building blocks.

The portfolio returns are often assumed to be normally distributed\(^{28}\). One of methods employing the normality assumption for returns is the \textit{delta} method (the delta-normal or the variance-covariance method).

\(^{28}\) JP Morgan, 1995; Phelan, 1995
2.1.3. Delta Method

The *delta* method estimates changes in prices of securities using their “deltas” with respect to basic risk factors. The method involves a *linear* (also named as *delta* or *local*) approximation to (log) price movements:

\[ P(X+U) = P(X) + P'(X) U, \]

or

\[ \Delta P(X) = P(X+U) - P(X) \approx P'(X) U, \]

where \( X \) is the level of the basic risk factor (i.e., an equity, an exchange rate), \( U \) is the change in \( X \), \( P(X+U) = P(t+\tau, X+U) \), \( P(X) = P(t, X) \), \( P(X) \) is the (log) price of the asset at the \( X \) level of the underlying risk factor, \( P'(X) = \partial P/\partial X \) is the first derivative of \( P(X) \), it is commonly called the *delta* \((\Delta = \Delta(X))\) of the asset.

Thus, the price movements of the securities approximately are

\[ \Delta P(X) \approx P'(X)U = \Delta U. \]

The *delta-normal* (the *variance-covariance*) method computes the portfolio VAR as

\[ \text{VAR} = -Y_d \sigma_{z_{-a}} \sqrt{d^T \Sigma d}, \]

where \( d = d(X) = (\Delta_1(X), \Delta_2(X), ..., \Delta_n(X))^T \) is a vector of the delta-positions, \( \Delta(X) \) is the security’s delta with respect to the j-th risk factor, \( \Delta_j = \partial P/\partial X_j \).

2.1.4. VAR Based on the Gamma Distribution Assumption

Since the normal model for factor distributions is overly simplistic, Fong and Vasicek (1997) suggest to estimate the probability distribution of the portfolio value changes \( \Delta P \) by another type of the parametric distributions - the *gamma distribution*. They also assume that the basic risk factors \( X_i \) are jointly normally distributed with the zero mean and the covariance matrix \( \Sigma \). However, Fong and Vasicek propose a *quadratic* (gamma or *delta-gamma*) approximation to the individual asset price changes:

---

29 Because the time horizon \((\tau)\) is fixed and \( t \) is the present time, we shall omit the time argument and shall write \( P(X+U) \) instead of underlying \( P(t+\tau, X+U) \) and \( P(X) \) instead of \( P(t, X) \). We shall consider the dependency of \( P \) on the risk factor \( X \) only.
\[ \Delta P(X) = P(X + U) - P(X) = \sum_{i=1}^{n} \Delta_i U_i + \sum_{i=1}^{n} \sum_{j=1}^{n} \Gamma_{ij} w_i w_j U_i U_j, \]

where \( \Delta P \) is a security price change, \( n \) is the number of basic risk factors, \( U_j \) is the change in the value of the \( j \)-th risk factor, \( \Delta_j = \Delta_j(X) \) is the security’s delta at the level \( X \) with respect to the \( j \)-th risk factor, \( \Delta_j = \partial P / \partial X_j \), \( \Gamma_j \) is quadratic exposure (the gamma) at the level \( X \) to the \( j \)-th risk factor, \( \Gamma_j = \Gamma_j(X) = \partial^2 P / \partial X_j^2 \), \( j = 1, \ldots, n \).

The delta-gamma approximation for the portfolio return in one period is defined by

\[
\Delta P = \Delta P(X) = P(X + U) - P(X) = \sum_{i=1}^{n} \Delta_i U_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \Gamma_{ij} w_i w_j U_i U_j, \tag{2}
\]

where \( X = (X_1, X_2, \ldots, X_n)^T \), \( X_i \) is \( i \)-th risk factor, \( U_i \) is the change in the risk factor \( X_i \), \( w_i \) is the weight of the \( i \)-th risk factor, \( \Gamma_{ij} = \Gamma_{ij}(X) \) is the portfolio \( (i,j) \)-gamma,

\[
\Gamma_{ij}(X) = \partial^2 P(X) / \partial X_i \partial X_j, \quad \Gamma_{jj} = \Gamma_j, \quad i = 1, \ldots, n, \quad j = 1, \ldots, n.
\]

The variance of portfolio return can be estimated by

\[
V(\Delta P(X)) = \sum_{i=1}^{n} \sum_{j=1}^{n} \Delta_i \Delta_j w_i w_j \text{cov}(X_i, X_j) + \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \Delta_i \Gamma_{jk} w_i w_j w_k \text{cov}(X_i, X_j, X_k) + \frac{1}{4} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \Gamma_{ij} \Gamma_{kl} w_i w_j w_k w_l \text{cov}(X_i, X_j, X_k, X_l).
\]

From (2), \( \Delta P \) is a quadratic function of normal variates. This distribution of \( \Delta P \) is, in general, non-symmetric. However, one can approximate the quantile by the skewness parameter and the standard deviation. In fact, Fong and Vasicek (1997) used the approximation for the portfolio VAR value, based on a generalized "gamma" distribution:

\[
\text{VAR} = -Y_0 k(\gamma, \alpha) \sigma(X_p),
\]

where \( \gamma \) is the skewness of the distribution, \( \gamma = \mu_3 / \sigma^3 \), \( \mu_3 \) is the third moment of \( \Delta P \), \( k(\gamma, \alpha) \) is the ordinate obtained from the generalized gamma distribution for the skewness \( \gamma \) at the confidence level \( \alpha \). Fong and Vasicek (1997) report the \( k(\gamma, \alpha) \) values at \( \alpha = 99 \% \):
<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$k(\gamma, \alpha)$</th>
<th>$\gamma$</th>
<th>$k(\gamma, \alpha)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2.83</td>
<td>3.99</td>
<td>0.50</td>
<td>1.96</td>
</tr>
<tr>
<td>-2.00</td>
<td>3.61</td>
<td>0.67</td>
<td>1.83</td>
</tr>
<tr>
<td>-1.00</td>
<td>3.03</td>
<td>1.00</td>
<td>1.59</td>
</tr>
<tr>
<td>-0.67</td>
<td>2.80</td>
<td>2.00</td>
<td>0.99</td>
</tr>
<tr>
<td>-0.50</td>
<td>2.69</td>
<td>2.83</td>
<td>0.71</td>
</tr>
</tbody>
</table>

Source: Fong and Vasicek, 1997

The gamma distribution takes into consideration the skewness of the $\Delta P$ distribution, whereas the normal distribution is symmetric and does not reflect the skewness.

2.2. Historical Simulation

The historical simulation approach constructs the distribution of the portfolio value changes $\Delta P$ from historical data without imposing distribution assumptions and estimating parameters. Hence, sometimes the historical simulation method is called a nonparametric method. The method assumes that trends of past price changes will continue in the future. Hypothetical future prices for time $t+s$ are obtained by applying historical price movements to the current (log) prices:

$$P_{i,t+s}^* = P_{i,t+s-1,t}^* + \Delta P_{i,t+s-\kappa,t}$$

where $t$ is the current time, $s = 1, 2, \ldots, \kappa$, $\kappa$ is the horizon length of going back in time, $P_{i,t+s}$ is the hypothetical (log) price of the $i$-th asset at time $t+s$, $P_{i,t}^* = P_{i,t}$, $\Delta P_{i,t+s-\kappa} = P_{i,t+s-\kappa} - P_{i,t+s-\kappa-1}$, $P_{i,t}$ is the historical (log) price of the $i$-th asset at time $t$. Here we assumed that the time horizon $\tau = 1$.

A portfolio value $P_{p,t+s}^*$ is computed using the hypothetical (log) prices $P_{i,t+s}^*$ and the current portfolio composition. The portfolio return at time $t+s$ is defined as

$$R_{p,t+s}^* = P_{p,t+s}^* - P_{p,t}$$

where $P_{p,t}$ is the current portfolio (log) price.
The portfolio VAR is obtained from the density function of computed hypothetical returns. Formally, \( \text{VAR} = \text{VAR}_{t,T} \) is estimated by the \((1-\alpha)\)th quantile, \( \text{VAR}^* \); namely, 
\[
F_{\kappa,\Delta \kappa}(-\text{VAR}^*) = 1 - \alpha, \quad \text{where} \quad F_{\kappa,\Delta \kappa}(x) \quad \text{is the empirical density function}
\]
\[
F_{\kappa,\Delta \kappa}(x) = \frac{1}{\kappa} \sum_{\kappa=1}^{\kappa} I\{R_{\kappa,\Delta \kappa} \leq x\}, \quad x \in \mathbb{R}.
\]

2.3 Monte Carlo Simulation

The Monte Carlo method specifies statistical models for basic risk factors and underlying assets. The method simulates the behavior of risk factors and asset prices by generating random price paths. Monte Carlo simulations provide possible portfolio values on a given date \( T \) after the present time \( t \), \( T > t \). The VAR \( \text{VAR}_T \) value can be determined from the distribution of simulated portfolio values. The Monte Carlo approach is performed according to the following algorithm:

1. Specify stochastic processes and process parameters for financial variables.
2. Simulate the hypothetical price trajectories for all variables of interest. Hypothetical price changes are obtained by simulations, draws from the specified distribution.
3. Obtain asset prices at time \( T \), \( P_{i,T} \), from the simulated price trajectories. Compute the portfolio value \( P_{p,T} = \sum w_{i,T} P_{i,T} \).
4. Repeat steps 2 and 3 many times to form the distribution of the portfolio value \( P_{p,T} \).
5. Measure \( \text{VAR}_T \) as the \((1-\alpha)\)th percentile of the simulated distribution for \( P_{p,T} \).

2.4 Stress testing

The parametric, historical simulation, and Monte Carlo methods estimate the VAR (expected losses) depending on risk factors when markets function without abrupt changes. The stress testing method examines the effects of large movements in key
financial variables on the portfolio value. The price movements are simulated in line with the certain scenarios\textsuperscript{30}. Portfolio assets are reevaluated under each scenario. The portfolio return is derived as

\[ R_{p,s} = \sum w_{i,s} R_{i,s}, \]

where \( R_{i,s} \) is the hypothetical return on the i-th security under the new scenario \( s \), \( R_{p,s} \) is the hypothetical return on the portfolio under the new scenario \( s \).

Estimating a probability for each scenario \( s \) allows to construct a distribution of portfolio returns, from which VAR can be derived.

3. Components of VAR Methodologies

Implementation of the VAR methodologies requires analysis of their components:

- distribution and correlation assumptions,
- volatility and covariance models,
- weighting schemes,
- the window length of data used for parameter estimations,
- the effect of the time horizon (holding period) on the VAR values, and
- incorporation of the mean of returns in the VAR analysis.

3.1. Distribution assumptions

The parametric VAR methods assume that asset returns have parametric distributions. The parametric approaches are subject to “model risk”: distribution assumptions might be incorrect. The frequent assumption is that asset returns have a multivariate normal distribution. Though, many financial time-series violate the normality assumption. Empirical data exhibit asymmetric, leptokurtic or platokurtic distributions with heavy

\textsuperscript{30} Scenarios include possible movements of the yield curve, changes in exchange rates, etc. together with estimates of the underlying probabilities.
tails. Fong and Vasicek (1997) suggest to use the gamma distribution. Analysts often conjecture assets are independent over time but correlated across assets. The simulation technique does not place the distribution assumptions, thus, it is free of model risk and “parameter estimation” risk. The Monte Carlo approach specifies the distributions of the underlying instruments.

3.2. Volatility and covariance models

The VAR methods apply diverse volatility and correlation models:

- constant volatility (moving window),
- exponential weighting,
- GARCH,
- EGARCH (asymmetric volatility),
- cross-market GARCH,
- implied volatility,
- subjective views

3.2.1. Constant Volatility Models

In the constant volatility (equally weighted) models, variances and covariances do not change over time. They are approximated by sample variances and covariances over the estimation “window”:

$$\hat{\sigma}_{t,T}^2 = \frac{1}{T-t} \sum_{t=i}^{T} (R_i - \hat{\mu}_{t,T})^2,$$

where $\hat{\sigma}_{t,T}$ is the estimated variance of returns $R_i$ over the time window $[t,T]$, $\hat{\mu}_{t,T}$ is the estimated mean of returns over the time window $[t,T]$.

---

31 Duffie and Pan, 1997; Jackson, Maude, and Perraudin, 1997; JP Morgan, 1995; Phelan, 1995; Hopper, 1997; Mahoney, 1995; Hendricks, 1996
\[ \hat{\mu}_{t,T} = \frac{1}{T-t} \sum_{i=t+1}^{T} R_i. \]

If the mean return is assumed to be sufficiently small,
\[ \hat{\sigma}_{t,T}^2 = \frac{1}{T-t} \sum_{i=t+1}^{T} R_i^2. \]

### 3.2.2. Weighted Volatility Models

The empirical financial data do not exhibit constant volatility. The *exponential weighting* models take into account time-varying volatility and accentuate the recent observations:

\[ \hat{\sigma}_{t,T}^2 = \frac{1}{T-t} \sum_{i=t+1}^{T} \theta_i (R_i - \hat{\mu}_{t,T})^2, \]

where \( \theta_i \) are the weighting values:

\[ 0 < \theta_i < 1, \quad \sum_{i=t+1}^{T} \theta_i = 1. \] (3)

The weighting schemes are divided on *uniform*\(^{32}\) and *asset-specific*\(^{33}\) schemes. The JP Morgan’s RiskMetrics system adopted the *uniform weighting* approach:

\[ \theta_i = (1-\lambda)\lambda^{c_{T-t}}, \]

where \( \lambda \) is the decay factor, \( 0 < \lambda < 1 \), and \( c_{T-t} \) > 0 is chosen so that the constraints (3) are met. JP Morgan uses \( \lambda = 0.94 \) for a 1-day time horizon.

Jackson, Maude and Perraudin (1997) demonstrate that the weighting schemes with lower values of \( \lambda \) in parametric models lead to higher tail probabilities, proportions of actual observations exceeding the VAR predictions\(^{34}\). They point out a trade-off between the degree of approximating time-varying volatilities and the performance of the parametric methods. Hendricks (1996) found that decreasing \( \lambda \) accompanies with higher variability of the VAR measurements.

---

\(^{32}\) JP Morgan, 1995; Jackson, Maude, and Perraudin, 1997

\(^{33}\) Lawrence and Robinson, 1995

\(^{34}\) see Jackson, Maude, Perraudin, 1997, table 4, p. 179
The CSFB’s PrimeRisk employs the asset-specific weighting schemes. It develops specific volatility models (different weighting schemes) for different types of assets (i.e., equities, futures, OTC options).

3.2.3. ARCH Models

Popular models explaining time-varying volatility are autoregressive conditional heterokedasticity (ARCH) models, introduced by Engle (1982). In the ARCH models the conditional variances follow autoregressive processes. The ARCH(q) model assumes the returns on the i-th asset \( R_{i,1}, R_{i,2}, \ldots \) are explained by the process:

\[
R_{i,t} = \mu_i + \sigma_{i,t}u_{i,t},
\]

\[
\sigma_{i,t}^2 = \alpha_i + \sum_{j=1}^{q} \beta_{ij}(R_{i,t-j} - \mu_i)^2,
\]

where \( \mu_i \) is the expectation of \( R_i \), \( \sigma_{i,t}^2 \) is the conditional variance of \( R_i \) at time \( t \), \( u_{i,t+1} \) is a random shock with the mean of zero and the variance of 1 (a common assumption is \((U_{i,t})_{t \geq 1} \sim iid \, N(0,1))\), \( \alpha_i \) and \( \beta_{ij} \) are constants, \( \alpha_i > 0, \beta_{ij} \geq 0, j=1,\ldots, q, i=1,\ldots, n \).

In the ARCH(1) model the conditional volatility at period \( t \) depends on the volatility at the previous period \( t-1 \). If volatility at time \( t-1 \) was large, the volatility at time \( t \) is expected to be large as well. Observations will exhibit clustered volatilities: one can distinguish periods with high volatilities and tranquil periods.

3.2.4. GARCH Models

Bollerslev (1986) suggested the generalized ARCH (GARCH) model\(^{36}\). In the GARCH models the conditional variance contains both autoregressive and moving average components (it follows an ARMA process). In the GARCH\((p,q)\) model, the return on the i-th asset has the representation

\[
R_{i,t} = \mu_i + \sigma_{i,t}u_{i,t},
\]

\(^{35}\) The dependence structure of the returns \( R_t = (R_{1,t}, R_{2,t}, \ldots, R_{n,t}) \) needs additional specifications for each \( t>0 \). See, for example, section 3.2.6 further on.

\(^{36}\) Bollerslev, 1986; Bollerslev, Chou, and Kroner, 1992
the conditional variance is assumed to follow

$$\sigma^2_{i,t} = \alpha_i + \sum_{j=1}^{q} \beta_j (R_{i,t-j} - \mu_i)^2 + \sum_{k=1}^{p} \gamma_k \sigma^2_{i,t-k},$$

where $\alpha_i$, $\beta_j$, $\gamma_k$ are constants, $\alpha_i > 0$, $\beta_{ij} \geq 0$, $\gamma_{ik} \geq 0$, $j=1,...,q$, $k=1,...,p$, $i=1,...,n$.

The advantage of using the GARCH model ensues from the fact that an AR process of an high order might be represented by more parsimonious ARMA process. Thus, the GARCH model will have less parameters to be estimated than the corresponding ARCH model.

3.2.5. EGARCH Models

Nelson (1991) introduced the exponential GARCH (EGARCH) model. In the general EGARCH($p$, $q$) model, the conditional variance follows\(^{37}\)

$$\log \sigma^2_t = \alpha + \sum_{j=1}^{q} \beta_j \frac{|R_{t-j} - \mu|}{\sigma_{t-j}} - E \frac{|R_{t-j} - \mu|}{\sigma_{t-j}} + \delta \frac{(R_{t-j} - \mu)}{\sigma_{t-j}} + \sum_{i=1}^{p} \gamma_i \log \sigma^2_{t-i}.$$

The $\delta$ parameter helps explain asymmetric volatility. If $\beta_j > 0$ and $-1 < \delta < 0$, then negative deviations of $R_t$ from the mean entail higher volatility than positive deviations do. If $\beta_j > 0$ and $\delta < -1$, then positive deviations lower volatility whereas negative deviations cause additional volatility.

The advantage of using the EGARCH model is that it does not impose positivity restrictions on coefficients, whereas the GARCH model requires coefficients to be positive.

\(^{37}\) Hamilton, 1994
3.2.6. Cross-market GARCH

The cross-market GARCH allows to estimate volatility in one market from volatilities in other markets. Duffie and Pan (1997) provide an example of cross-market GARCH, which employs the bivariate GARCH model:

\[
\begin{bmatrix}
\sigma^2_{1,t} \\
\sigma^2_{2,t} \\
\sigma_{12,t}
\end{bmatrix} = A + B \begin{bmatrix}
R^2_{1,t} \\
R_{1,t}R_{2,t} \\
R^2_{2,t}
\end{bmatrix} + \Gamma \begin{bmatrix}
\sigma^2_{1,t-1} \\
\sigma^2_{2,t-1} \\
\sigma_{12,t-1}
\end{bmatrix}
\]

where \( \sigma_{1,t-1} \) is the conditional standard deviation of \( R_{1,t} \), \( \sigma_{2,t-1} \) is the conditional standard deviation of \( R_{2,t} \), \( \sigma_{12,t-1} \) is the conditional covariance between \( R_{1,t} \) and \( R_{2,t} \), \( R_{1,t} \) is the return in the first market at time \( t \), \( R_{2,t} \) is the return in the second market at time \( t \), \( A \) is a vector of three elements, \( B \) is a 3x3 matrix, \( \Gamma \) is a 3x3 matrix.

3.2.7. Implied Volatilities

Sometimes analysts use implied volatilities to estimate future volatilities. Implied volatilities are volatilities derived from pricing models. For instance, implied volatilities can be obtained from the Black-Scholes option pricing model. Option prices calculated by the Black-Scholes formula \( C_t = C(S_t, K, r, \sigma, \tau) \) are increasing in volatility \( \sigma \). Hence, “inverting” the formula, one can obtain the implied volatility values \( \sigma = \sigma(C_t, S_t, K, r, \tau) \). Here, \( C_t \) is the option price, \( S_t \) is the price of the underlying asset, \( K \) is the exercise price, \( r \) is the constant interest rate, and \( \tau \) is the time to expiration.

The implied tree technique\(^{38}\) assumes implied volatilities change over time and computes them relating the modeled and observed option prices.

One of methods for estimating volatility is the method of "subjective views"\(^{39}\). Analysts make predictions of volatility from own views of market conditions.

Besides the distribution assumptions and volatility models, the VAR computations also need specification of correlation assumptions on price changes and volatilities within

\(^{38}\) Derman and Kani, 1994; Rubinstein, 1994; Jackwerth and Rubinstein, 1996

\(^{39}\) Hopper, 1997
and across markets\textsuperscript{40}. Beder (1995) illustrated the sensitivity of VAR results to correlation assumptions. She computed VAR using the Monte Carlo simulation method under different assumptions: (i) correlations across asset groups and (ii) correlations only within asset groups. The obtained VAR estimates were lower for the first type of correlation assumptions than for the second type.

3.3. Time horizon

The time horizon (the holding period) in the VAR computations can take any time value. In practice, it varies from one day to two weeks (10 trading days) and depends on liquidity of assets and frequency of trading transactions. It is assumed that the portfolio composition remains the same over the holding period. This assumption constrains dynamic trading strategies. The Basle Committee recommends to use the 10-day holding period. Users argue that the time horizon of 10 days is inadequate for frequently traded instruments and is restrictive for illiquid assets. Long holding periods are usually recommended for portfolios with illiquid instruments. Though, many model approximations are only valid within short periods of time.

Beder (1995) analyzed the impact of the time horizon on VAR estimations. She calculated VAR for three hypothetical portfolios applying four different approaches for the time horizons of 1-day and 10-day. For all VAR calculations, with the exception of one case, Beder reported larger VAR estimates for longer time horizons.

3.4 Window length

The window length is the length of the data subsample (the observation period) used for a VAR estimation. The window length choice relates to sampling issues and availability of databases. The regulators suggest to use the 250-day (one-year) window length.

Jackson, Maude, and Perraudin (1997) computed parametric and simulation VARs for the 1-day and 10-day time horizons using the window lengths from three to 24 months.

\textsuperscript{40} Duffie and Pan, 1997; Beder, 1995; Liu, 1996
They concluded that VAR forecasts based on longer data windows are more reliable\textsuperscript{41}. Beder (1995) estimated VAR applying the historical simulation method for the 100-day and 250-day window lengths. Beder shows that the VAR values increase with the expanded observation intervals. Hendricks (1996) calculated the VAR measures using the parametric approach with equally weighted volatility models and the historical simulation approach for window lengths of 50, 125, 250, 500, and 1250 days. He reports that the VAR measures become more stable for longer observation periods.

3.5. Incorporation of the mean of returns

In many cases the mean of returns is assumed to be zero. Jackson, Maude, and Perraudin (1997) analyze the effects of (i) inclusion of the mean in calculations and (ii) setting the mean to zero on VAR results. Their analysis did not lead to certain conclusions\textsuperscript{42}.

4. Evaluation of VAR Methods

The VAR methods provide only estimated VAR values. The important problem is to evaluate accuracy of the VAR estimates. Researchers propose different performance measures.

Kupiec (1995) compares techniques for assessing the accuracy of the VAR methods. He describes daily monitoring tests based on: the time until the first failure, the time between additional failures, the number of failures; a test for infrequent monitoring schemes based on the sample proportion test statistics and the binomial distribution; tests based on historical simulation. Kupiec concludes: the tests based on observed data have low power; the critical values, obtained from historical simulations, are biased and subject to sampling errors. Lopez (1996), and Mahoney (1995) also investigate evaluations based on the binomial distribution. Hendricks and Hirtle (1997) explain back-testing of VAR results using the BIS red, yellow, and green frequency zones. Crnkovic

\textsuperscript{41} Jackson, Maude, and Perraudin, 1997, table 5, p. 180
\textsuperscript{42} Jackson, Maude, and Perraudin, 1997, table 6, p. 181

Hendricks (1996) evaluates performance of the twelve VAR approaches for foreign exchange portfolios using nine criteria: mean relative bias, root mean squared relative bias, annualized percentage volatility, fraction of outcomes covered, multiple needed to attain desired coverage, average multiple of tail event to risk measure, maximum multiple of tail event to risk measure, correlation between risk measure and absolute value of outcome, and mean relative bias for risk measures scaled to desired level of coverage. Hendricks infers that almost all considered methodologies give the accurate VAR estimates at the 95 percent confidence level. At the 99 percent confidence level, the performance of the VAR methods is not sound. The historical simulation VAR estimates tend to be greater than the variance-covariance VAR estimates based on the normality assumptions. His results confirm that, in practice, extreme movements are more frequent than it is assumed by the normal distribution.

The trade-offs among accuracy, speed, and cost are considered by Pritsker (1997) and Robinson (1997).

5. VAR Strengths and Weaknesses

The VAR methodologies are becoming necessary tools in risk management. It is important to be aware of VAR strengths and weaknesses. Institutions use the VAR measurements to estimate exposure to market risks and

---

43 Beder, 1995, 1997; Mahoney, 1995; Simons, 1996; Jorion, 1996, 1997; Hopper, 1997; Shaw, 1997;
assess expected losses. Application of different VAR methods provides different VAR estimates. The choice of methods should mostly depend on the portfolio composition. If a portfolio contains instruments with linear dependence on basic risk factors, the delta method will be satisfactory. Strength of the delta approach is that computations of VAR are relatively easy. Drawbacks of the delta-normal method are: (i) empirical observations on returns of financial instruments do not exhibit the normal distribution and, thus, the delta-normal technique does not fit well data with heavy tails; (ii) accuracy of VAR estimates diminishes with nonlinear instruments: in their presence, VAR estimates are understated. For portfolios with option instruments, historical and Monte Carlo simulations are more suitable. The historical simulation method is easy to implement having sufficient database. The advantage of using the historical simulation is that it does not impose distributional assumptions. Models based on historical data assume that the past trends will continue in the future. However, the future might encounter extreme events. The historical simulation technique is limited in forecasting the range of portfolio value changes. The stress-testing method can be applied to investigate effects of large movements in financial variables. A weakness of stress-testing is that it is subjective. The Monte Carlo method can incorporate nonlinear positions and non-normal distributions restricted. It does not restrict the range of portfolio value changes. The Monte Carlo method can be used in conducting the sensitivity analysis. The main limitations in implementing the Monte Carlo methodology are: (i) it is affected by model risk; (ii) computations and software are complex; (iii) it is time consuming.

VAR methodologies are subject to implementation risk: implementation of the same model by different users produces different VAR estimates. Marshall and Siegel (1997) conducted an innovative study of implementation risk. They compared VAR results obtained by several risk management systems developers using one model, JP Morgan’s RiskMetrics. Marshal and Siegel found that, indeed, different systems do not produce the same VAR estimates for the same model and identical portfolios. The varying estimates can be explained by the sensitivity of VAR models to users’ assumptions. The degree of variation in VAR numbers was associated with the portfolio composition. Dependence of implementation risk on instrument complexity can be summarized in the following

Schachter, 1997
relative ascending ranking: foreign exchange forwards, money markets, forward rate agreements, government bonds, interest rate swaps, foreign exchange options, and interest rate options. Nonlinear securities entail larger discrepancy in VAR results than linear securities. In order to take into account implementation risk, it is advisable to accompany VAR computations for nonlinear portfolios with sensitivity analysis to underlying assumptions.

Other VAR weaknesses are:

• Existing VAR models reflect observed risks and they are not useful in transition periods characterized by structural changes, additional risks, contracted liquidity of assets, and broken correlations across assets and across markets.

• The trading positions change over time. Therefore, extrapolation of a VAR for a certain time horizon to longer time periods might be problematic. Duffie and Pan (1997) point out that if intra-period position size is stochastic, then the VAR measure obtained under the assumption of constant position sizes, should be multiplied by a certain factor\(^44\).

• The VAR methodologies assume that necessary database is available. For certain securities, data over a sufficient time interval may not exist.

• If historical information on financial instruments is not available, the instruments are mapped into known instruments. Though, mapping reduces precision of VAR estimations.

• Model risks can occur if the chosen stochastic underlying processes for valuing securities are incorrect.

• Since true parameters are not observable, estimates of parameters are obtained from sample data. The measurement error rises with the number of estimated parameters.

• VAR is not effective when strategic risks are significant.

\(^{44}\) Duffie and Pan (1997) provide an expression for the factor in the case of a single asset. If: (i) the underlying asset returns have constant volatility \(\sigma\), (ii) the position size is a martingale and follows lognormal process with volatility \(s\), then a multiplication factor is approximately of \((\exp(at) -1)/(at)\), where \(a=2s^2+4ps\sigma\), \(p\) is the correlation of the position size with the asset.
6. VAR Improvements

In order to improve performance of VAR methodologies, researchers suggest numerous modifications of traditional techniques and new ways of VAR evaluations. This section presents modifications of:

- delta,
- historical simulation,
- Monte Carlo,
- scenario analysis.

The section also describes new approaches to VAR estimations and interpretations.

6.1. Modifications of Delta

Let \( U_t \) be the \( n \)-dimensional vector of changes in risk factors over one period, \( U_t = \Delta X_t \). The standard delta and delta-gamma methods assume that changes in risk factors follow the normal distribution conditional on the current information:

\[
(U_{t+1} | \Omega_t) \sim \mathcal{N}(0, \Sigma_t),
\]

where \( \Omega_t \) is the information available until the current time \( t \).

Delta methods apply a linear approximation to the portfolio returns as a function of the underlying risk factors:

\[
\Delta P_{t+1} \equiv P_{t+1}(X_t + U_t) - P_{t+1}(X_t) = \Delta_t^T Y_{t+1},
\]

where \( \Delta_t = (\Delta_{1,t}, \Delta_{2,t}, ..., \Delta_{n,t})^T, \Delta_t = \Delta_t(X_t) = \frac{\partial P_t(X_t)}{\partial X_t}, X_t \) is the vector of risk factors, \( X_t = (X_{1,t}, X_{2,t}, ..., X_{n,t})^T, P_t(X_t) \) is the portfolio log-price at time \( t \) which depends on the current risk factors only, and \( Y_{t+1} = (Y_{1,t+1}, Y_{2,t+1}, ..., Y_{n,t+1})^T \), where \( Y_{i,t} = w_i U_{i,t}, i=1, ..., n \).

Under the delta approach, \( \Delta P_{t+1} \sim \mathcal{N}(0, \Delta_t^T \Sigma_Y \Delta_t) \), where \( \Sigma_Y \) is the covariance matrix of \( Y_{t+1} \).

\[45\] Formally, \( (\Omega_t)_{\geq 0} \) is the filtration generated by the underlying market-shot-process (processes), typically
**Delta-gamma methods** use a quadratic approximation to the portfolio value changes:

\[ \Delta P_{t+1} \approx \Delta_t^T Y_{t+1} + \frac{1}{2} Y_{t+1}^T \Gamma_t Y_{t+1} \]

where

\[ \Gamma_t = \Gamma(t, X_t) = \frac{\partial^3 P_t(X_t)}{\partial X_t \partial X_t^T}. \]

Under *delta-gamma methods*, the distribution of \( \Delta P \) cannot be approximated by the normal distribution. Hence, a traditional technique of VAR derivation as a multiple of the \((1-\alpha)\)-th percentile of the standard normal distribution cannot be employed.

We shall describe the following improvements of the delta-gamma method:

- Delta-gamma-Monte Carlo,
- Delta-gamma-delta,
- Delta-gamma-minimization,
- Delta-gamma-Johnson,
- Delta-gamma-Cornish-Fisher.

### 6.1.1. Delta-gamma-Monte Carlo

The *delta-gamma-Monte Carlo* method approximates the distribution of \( \Delta P \) by the distribution of hypothetical portfolio value changes\(^{46}\):

(i) values of \( U_{t+1} \) are obtained by random draws from its distribution. Pritsker (1997) mentions two types of draws: historical simulation and bootstrapping;

(ii) hypothetical values of \( \Delta P \) are calculated at each draw using the delta-gamma approximation;

(iii) steps (i) and (ii) are repeated many times;

(iv) the distribution of \( \Delta P \) is formed by ordering the \( \Delta P \) values from step (iii).

\(^{46}\) Pritsker, 1996/1997
A VAR estimate is derived as the \((1-\alpha)\)-th percentile of the \(\Delta P\) distribution constructed in step (iv).

### 6.1.2. Delta-gamma-delta

The *delta-gamma-delta* method also employs a delta-gamma approximation\(^{47}\). It assumes: (i) shocks to the portfolio value \(P\) are represented by \(U_{t+1}\) and elements of \(U_{t+1}U_{t+1}'\), which correspond to \(\Delta\) and \(\Gamma\) terms in a delta-gamma approximation; (ii) the shocks are uncorrelated and normally distributed. For instance, if the portfolio value depends on a single factor, then shocks to \(P\) are assumed to come from jointly normally distributed \(U_{t+1}\) and \(U_{t+1}^2\). Though, the assumptions of the normality \(U_{t+1}^2\) and the joint normality are not correct. According to the delta-gamma-delta approach,

\[
\Delta P_{t+1} \sim \mathcal{N}(0.5 \Gamma_t \sigma_t^2, \Delta_t^2 \sigma_t^2 + 0.5 \Gamma_t^2 \Delta_t^4),
\]

where \(\sigma_t^2\) is the variance of \(Y_{t+1}\), \(Y_{i,t+1} = w_i U_{i,t+1}\). Therefore, \(VAR\) can be calculated as

\[
VAR_t = \frac{1}{2} \Gamma_t \sigma_t^2 + z_{1-\alpha} \sqrt{\Delta_t^2 \sigma_t^2 + \frac{1}{2} \Gamma_t^2 \sigma_t^4}.
\]

### 6.1.3. Delta-gamma-minimization

The *delta-gamma-minimization* method uses a delta-gamma approximation to \(\Delta P\) and assumes that a vector of changes in risk factors \(U_{t+1}\) is normally distributed\(^{48}\). Denote by \(Y_{t+1}\) a vector of shocks with the weights of risk factors, \(Y_{i,t+1} = w_i U_{i,t+1}\). The delta-gamma-minimization technique determines \(VAR_t\) as the solution of the following minimization problem:

\[
VAR_t = \min_{Y_{t+1}} \left\{ \Delta_t^T Y_{t+1} + \frac{1}{2} Y_{t+1}^T \Gamma_t Y_{t+1} \right\}
\]

subject to the constraint

\[
\Delta_t^T \mathbf{I} + \frac{1}{2} \Gamma_t^2 \mathbf{I} = \alpha - \mathbf{1}
\]

\[^{47}\text{Pritsker, 1996/1997}\]

\[^{48}\text{Pritsker, 1996/1997; Fallon, 1996; Wilson, 1994}\]
\[ Y_{t+1}^T \Sigma^{-1} Y_{t+1} \leq \chi^2(\alpha, k), \]

where \( \Sigma \) is the covariance matrix of risk factors and \( \chi^2(\alpha, k) \) is the \( \alpha \% \) critical value of the central chi-squared distribution with \( k \) degrees of freedom. The method implicitly supposes that the \( \Delta P \) values within the constraint set exceed the "external" \( \Delta P \) values. In practice, this assumption might be violated. Hence, the fraction of the \( \Delta P \) values, which are lower than VAR, can be less than \( (1-\alpha)\% \) and the VAR estimate can be overstated. The strengths of the delta-gamma-minimization method are: (i) it does not impose the assumption of joint normality as the delta-gamma-delta method does; (ii) it avoids Monte Carlo data generation.

6.1.4. Delta-gamma-Johnson

The \textit{delta-gamma-Johnson} method\(^{49}\) relies on the normal assumption for the distribution of \( (U_t)_{t \geq t} \). The method chooses a distribution function for \( \Delta P \) and estimates its parameters matching the first four moments of the distribution and the delta-gamma approximation of \( \Delta P \). A VAR estimate is obtained from the cumulative density function of the chosen distribution. The strength of the method is that it is analytic. However, the delta-gamma-Johnson method uses information only up to the fourth moment and might be less precise than the delta-gamma-Monte-Carlo method, which uses all information on the delta-gamma Taylor expansion.

6.1.5. Delta-gamma-Cornish-Fisher

The \textit{delta-gamma-Cornish-Fisher} method is based on a delta-gamma approximation of the \( \Delta P \) distribution and the normality assumption for \( U_t \). It uses a Cornish-Fisher expansion to estimate the \((1-\alpha)\)-th percentile of the standardized \( \Delta P \) distribution \( \Delta P^S \).\(^{50}\)

\(^{49}\) Zangari, 1996; Pritsker, 1996/1997
\(^{50}\) Fallon, 1996; Pritsker, 1996/1997; Zangari (1996)
where $q_3$ is the third cumulant of $\Delta P^S$ and $q_4$ is the fourth cumulant of $\Delta P^S$. The cumulants $q_i$ are determined from an expansion

$$\ln(G_{\Delta P^S}(t)) = \sum_{i=1}^{\infty} q_i \frac{t^i}{i!},$$

where $G_{\Delta P^S}(t) = E[\exp(t(\Delta P^S))]$ is the moment generating function of $\Delta P^S$.

The advantage of the delta-gamma-Cornish-Fisher approach is that it is analytic. The weakness of the method: it ignores part of information.

6.2. Modifications of Historical Simulation

The standard historical simulation technique forms the distribution of portfolio returns based on the historical data. It obtains future returns on assets by applying past changes to the current returns.

We consider here two modifications of the historical simulation method:

- bootstrapped historical simulation and
- combining kernel estimation with historical simulation.

6.2.1. Bootstrapped Historical Simulation

The *bootstrapped historical simulation* method generates returns of the risk factors by "bootstrapping" from historical observations. It is suggested to draw data from the updated returns. Suppose that we have observations of returns on $n$ assets $[R_{it}]_{1 \leq i \leq n, 1 \leq t \leq T}$ with the covariance matrix $\Sigma$. The past returns can be updated using new estimates of

(i) volatilities, and

---

51 Duffie and Pan, 1997
(ii) volatilities and correlations.

In case (i), the updated returns are given by

\[ R_{it}^U = R_{it} \frac{\sigma_i^U}{\sigma_i}, \]

where \( \sigma_i \) is the historical standard deviation (volatility) of the i-th asset, \( \sigma_i^U \) = a new estimate of the standard deviation \( \sigma_i \), \( R_{it}^U \) is an updated return for the i-th asset, \( t=1, \ldots, T \), \( i=1, \ldots, n \).

In case (ii), the updated return vector is determined by

\[ R_i^U = (\Sigma^U)^{\frac{1}{2}} (\Sigma^{-\frac{1}{2}} R_i, \]

where \( R_i = (R_{1i}, R_{2i}, \ldots, R_{ni})^T \) is the vector of returns at time \( t \), \( (\Sigma^{-\frac{1}{2}} \) is the matrix square root of \( \Sigma^{-1} \), and \( (\Sigma^U)^{\frac{1}{2}} \) is the matrix square root of the updated covariance matrix \( \Sigma^U \).

One of future research questions would be to investigate impacts of updating approaches on the shape of the portfolio return distribution.

6.2.2. Combining Kernel Estimation with Historical Simulation

Butler and Schachter (1996) suggest to combine the historical simulation approach with kernel estimation. This modification allows to estimate precision of VAR measures and to construct confidence intervals around them. The combined approach is performed in three steps:

1. Approximation of the probability density function (pdf) \( f_{\Delta P}(x) \) and the cumulative distribution function (cdf) \( F_{\Delta P}(x) \) of portfolio returns,

2. Approximation of the distribution of the order statistic corresponding to the confidence level.\(^{52} \)

\(^{52}\) Given a sample \( R_1, \ldots, R_k \) of observations \( R_i \), we rearrange the sample in increasing order \( R_{1,k} \leq \ldots \leq R_{k,k} \) (the random sample), then the \( i \)th order statistic is given by \( R_{i,k} \). The value of the \( q \)-quantile \((0 \leq q \leq 1)\) is the
3. Estimation of VAR using moments of the pdf for the $i$th order statistic $R_{i,s}$, determined by $(1-\alpha)$th quantile. The first moment, the mean, $ER_{i,k}$, approximates the VAR and the second moment, the variance, $\text{Var}R_{i,k}$ reflects precision of the VAR estimate. The standard deviation can be used for constructing confidence intervals.

**Step 1.** While the standard historical simulation technique derives the piecewise pdf, a kernel estimation forms a smooth pdf. Butler and Schachter (1996) apply a normal kernel estimator:

$$
\hat{f}_{\Delta P}(x) = \frac{1}{k(0.9\sigma)^{0.5}} \sum_{i=1}^{k} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x - R_i}{0.9\sigma}\right)^2},
$$

where $k$ is the sample size, $R_i$ is the $i$-th observation on portfolio return $\Delta P$, $i=1,...,k$, $\sigma$ is the standard deviation of $\Delta P$, $x$ is a current point of estimation.

The cdf $F_{\Delta P}(x)$ can be approximated:

- from the estimated pdf $\hat{f}_{\Delta P}(x)$, given by (4), or
- by the empirical cumulative distribution function

$$
\hat{F}_{\Delta P}(x) = \frac{1}{k} \sum_{i=1}^{k} 1\{X_i \leq x\}
$$

**Step 2.** Let $s$ be the $i$th order statistic, $h(s)$ the pdf of $s$, $H(s)$ the cdf of $s$. In order to assess the pdf $h(s)$ and the cdf $H(s)$, Butler and Schachter (1996) employ the estimated pdf $\hat{f}_{\Delta P}(x)$ from (4) and the empirical cdf $\hat{F}_{\Delta P}(x)$:

$$
\hat{H}(s) = \sum_{j=i}^{k} \frac{k!}{j!(k-j)!} \hat{F}_{\Delta P}(s)^j (1-\hat{F}_{\Delta P}(s))^{k-j},
$$

$$
\hat{h}(s) = \frac{k!}{i!(k-i)!} \hat{f}_{\Delta P}(x) \hat{F}_{\Delta P}(x)^{i-1} (1-\hat{F}_{\Delta P}(x))^{k-i}.
$$

---

value of the (rounded) $(qk)$th term in the ranked sample. Using historical return data $R_1, ..., R_k$, VAR can be estimated by the $(1-\alpha)$th quantile.

53 Butler and Schachter (1996) use the empirical cdf $\hat{F}_{\Delta P}(x)$ to approximate $F_{\Delta P}(x)$.

54 Note that $\hat{h}(s)$ is not equal to the derivative of $\hat{H}(s)$, $\hat{h}(s)$ is only "close" to $\frac{\partial \hat{H}(s)}{\partial s}$. 
Step 3. Moments of the pdf \( \hat{h}(s) \) can be obtained using 12-point Gauss-Hermite integration:

\[
E(s) = \int_{-\infty}^{\infty} s h(s) ds \approx \frac{1}{12} \sum_{j=1}^{12} a_j s_j e^{s_j^2} \hat{h}(s_j) = \hat{E}(s),
\]

\[
\text{Var}(s) = \int_{-\infty}^{\infty} s^2 h(s) ds - E(s)^2 \approx \frac{1}{12} \sum_{j=1}^{12} a_j s_j^2 e^{s_j^2} \hat{h}(s_j) - \hat{E}(s)^2 = \hat{\sigma}(s)^2,
\]

where \( a_j \) is the j-th integration weight, \( s_j \) is the j-th point of integral approximation, \( j = 1, \ldots, 12 \).

By the combined historical simulation and kernel estimation method, the VAR estimate is

\[
\text{VAR}^* = \frac{1}{12} \sum_{j=1}^{12} a_j s_j e^{s_j^2} \hat{h}(s_j). \tag{11}
\]

The \( \hat{\sigma}(s) \) can be used to form the confidence interval for \( \text{VAR}^* \) in large samples. If the sample is long enough, then the quantile is distributed normally. Therefore, for large samples\(^{55} \), the \( \beta \% \) confidence interval can be constructed as

\[
(\text{VAR}^* - z_{\beta} \hat{\sigma}(s), \text{VAR}^* + z_{\beta} \hat{\sigma}(s)),
\]

where \( z_{\beta} \) is the \( \beta \)-th quantile of the standard normal distribution.

6.3. Modifications of Monte Carlo

6.3.1. Quasi Monte Carlo

6.3.2. Grid Monte Carlo

6.4. Modifications of scenario analysis

6.4.1. Worst Case Scenario analysis

6.4.2. Factor-based interest rate scenarios

\(^{55} k > 50 \) is typically sufficient, but exact estimates for the threshold \( k \) are not known because in practice the financial returns form non-stationary processes.
6.4.3. Worst case distributional assumptions

6.5. DelVar. Evaluating effect of a new trade on VAR

6.6. Fitted probability distribution

6.7. Principal components and VAR

6.8. Spread risks within a VAR model

6.9. A confidence level on the profit and loss distribution

6.10. Overall profit and loss distribution

6.11. Maximum loss

6.12. Investor utility

6.13. A tail-fitting VAR model


6.15. VAR for blocks and VAR for a portfolio

6.16. A non-parametric VAR measure incorporating economic valuation

6.17. Minimization of VAR using options

6.18. Heavy tails

   6.18.1. Student t-distributions

   6.18.2. Mixture of normal densities

   6.18.3. Combining the mixture of normals and the quasi-Bayesian
            maximum likelihood estimation

6.19. Tail emphasized model optimization

6.20. Implied covariance forecasting

6.21. A second-order approximation to portfolio function and a multivariate
        GARCH(1,1) for state variables.

6.22. Quadratic normal distribution function for the change in portfolio value

**Part B. Future Research**

Despite the great importance of VAR, the existing methods do not provide a satisfactory evaluation of VAR. The main drawback is the lack of a convincing unified model for VAR capturing the following phenomena generally observed in financial data, such as asset returns, interest rates, exchange rates, equities:

- heavy tails of the marginal distributions of the process of financial returns,
- time-varying volatility,
- short- and long-range dependence.

The goal of this work is to construct models that encompass the above mentioned empirically observed features, and develop more precise VAR-estimation techniques based on the new models. A comparison with the existing methodologies will be given.

**7. VAR Models for Financial Returns with Heavy Tails**

The essence of VAR modeling is the prediction of the highest expected loss for a given portfolio. As we have seen in Part A, the VAR methods are designed to estimate low quantiles in the portfolio return distribution. The delta methods are based on the normal assumption for the distribution of financial returns. However, financial data violate the normality assumption. The empirical observations exhibit "fat" tails and excess kurtosis. The historical method does not impose the distributional assumptions but it is not reliable in estimating low quantiles with a small number of observations in tails. The performance of the Monte Carlo method depends on the quality of distributional assumptions of underlying risk factors.

Adequate approximation of distributional forms of returns is a key condition for accurate VAR estimation. We shall first analyze the issue of heavy tails. Given the leptokurtic nature (heavy tails and excess kurtosis) of empirical financial data, the stable Paretian distributions seem to be the most appropriate distributional models for asset
returns\textsuperscript{56}. We shall consider applications of stable distributions in VAR computations\textsuperscript{57}. First, we provide a finance-oriented description of stable distributions. Then, we describe modeling VAR with stable distributions.

7.1. A Finance-oriented Description of Stable Distributions

Definition 1: A random variable $R$ is said to be stable\textsuperscript{58} if for any $a > 0$ and $b > 0$ there exist constants $c > 0$ and $d \in \mathbb{R}$ such that

$$aR_1 + bR_2 \overset{d}{=} cR + d,$$

where $R_1$ and $R_2$ are independent copies of $R$ and $\overset{d}{=}$ denotes the equality in distribution.

A random variable $R$ is stable if and only if its characteristic function is described as

$$\Phi_R(\theta) = E(\exp(iR\theta)) = \exp\left\{-\sigma^\alpha |\theta|^\alpha \left(1 - i\beta (\text{sgn} \theta) \tan \frac{\pi \alpha}{2}\right) + i\mu \theta \right\}, \text{ if } \alpha \neq 1,$$

$$\Phi_R(\theta) = E(\exp(iR\theta)) = \exp\left\{-\sigma |\theta| \left(1 + i\frac{2}{\pi} (\text{sgn} \theta) \ln |\theta| \right) + i\mu \theta \right\}, \text{ if } \alpha = 1,$$

where $\alpha$ is the index of stability, $0 < \alpha \leq 2$, $\beta$ is the skewness parameter, $-1 \leq \beta \leq 1$, $\sigma$ is the scale parameter, $\sigma \geq 0$, and $\mu$ is the location parameter, $\mu \in \mathbb{R}$.

To indicate the dependence of a stable random variable $R$ from its parameters, we write $R \sim S_\alpha(\beta, \sigma, \mu)$.

In general, the stable distributions do not have closed form expressions for density and distribution functions.

If the index of stability $\alpha = 2$, then the stable distribution reduces to the Gaussian distribution. In empirical studies, typically, the modeling financial return data is done with stable distributions having $1 < \alpha < 2$ \textsuperscript{59}.


\textsuperscript{57} See also Gambrowski and Rachev, 1996.

\textsuperscript{58} Often $R$ is called $\alpha$-stable or Pareto stable (if $\alpha < 2$) or or Pareto-Lévy-stable (for $\alpha < 2$).

\textsuperscript{59} The financial returns modelled with $\alpha$-stable laws exhibit finite mean but infinite variance.
Stable distributions are unimodal and the smaller \( \alpha \) is, the stronger the leptokurtic feature of the distribution (the peak of the density becomes higher and the tails are heavier). Thus, the index of stability can be accepted as a measure of kurtosis.

In VAR estimations we are interested to investigate the behavior of the distributions in tails. The tails of the stable (non-Gaussian) distributions have power decay and are characterized by the following properties:

\[
\lim_{\lambda \to \infty} \lambda^\alpha P(R > \lambda) = k_\alpha \frac{1+\beta}{2} \sigma^\alpha
\]

and

\[
\lim_{\lambda \to \infty} \lambda^\alpha P(R < -\lambda) = k_\alpha \frac{1-\beta}{2} \sigma^\alpha,
\]

where

\[
k_\alpha = \begin{cases} 
1-\alpha & \text{if } \alpha \neq 1, \\
\frac{2}{\pi} & \text{if } \alpha = 1^{60}.
\end{cases}
\]

When \( \alpha > 1 \), the location parameter \( \mu \) measures the mean of the distribution.

If the skewness parameter \( \beta = 0 \), the distribution of \( R \) is symmetric and the characteristic function is

\[
\Phi_R(\theta) = \text{E}(\exp(iR\theta)) = \exp \left\{ -\sigma^\alpha |\theta|^\alpha + i\mu \theta \right\}
\]

If \( \beta > 0 \), the distribution is skewed to the right. If \( \beta < 0 \), the distribution is skewed to the left. Larger magnitudes of \( \beta \) indicate stronger skewness.

**Definition 2**: If \( \beta = 0 \) and \( \mu = 0 \), then the stable random variable \( R \) is called symmetric \( \alpha \)-stable (s\( \alpha \)s).

The scale parameter (the volatility) \( \sigma \) allows to write any stable random variable \( R \) as \( R = \sigma R_0 \), where \( R_0 \) has a unit scale parameter and the same index of stability \( \alpha \) and the same skewness parameter \( \beta \) as \( R \). The scale parameter generalizes the definition of the standard deviation. The stable analog of the variance is the variation: \( \nu_\alpha(R) = \sigma^\alpha(R) \).

---

60 Note that, in contrast to the normal case, the tails of the non-Gaussian (Pareto) stable distributions are much fatter, which will be an important issue in estimating VAR.
Definition 3: Absolute central moments of stable random variable with $\alpha > 1$ are defined for any positive real $p < \alpha$ as follows:

$$\|R\|_p = E\left( |R - E(R)|^p \right)^{1/p}.$$  

The $p$-th absolute moment, $E|R|^p$, is

- finite if $p < \alpha$ or $\alpha = 2$, and
- infinite otherwise.

The second moment of any non-Gaussian stable distribution is infinite.

Stable distributions possess the additivity property: a linear combination of stable random variables with the stability index $\alpha$ is a stable random variable with the same $\alpha^{61}$.

**Example:** If $R_1, R_2, \ldots, R_n$ are independent stable random variables with the stability index $\alpha$, $R_i \sim S_{\alpha}(\beta_i, \sigma_i, \mu_i)$, then $R = \sum_{i=1}^{n} R_i$ is a stable random variable with the same $\alpha$ and parameters

$$\sigma = \left(\sigma_1^\alpha + \cdots + \sigma_n^\alpha\right)^{1/\alpha}, \quad \beta = \frac{\beta_1 \sigma_1^\alpha + \cdots + \beta_n \sigma_n^\alpha}{\sigma_1^\alpha + \cdots + \sigma_n^\alpha}, \quad \mu = \mu_1 + \cdots + \mu_n.$$  

Suppose that two stable random variables $R_1$ and $R_2$ have the same index of stability $\alpha$, skewness parameter $\beta$, location parameter $\mu$; scale parameters $\sigma_1$ and $\sigma_2$, cumulative distribution functions $F_1$ and $F_2$, respectively. Then, these two distributions can be stochastically ordered of order 1 by their scale parameters, that is,

- $\sigma_1 \leq \sigma_2$ if and only if $\int_{-\infty}^{x} F_1(t)dt \geq \int_{-\infty}^{x} F_2(t)dt$, for all $x$; equivalently,
- the first distribution is preferred to the second by risk-averse agents if and only if $\sigma_1 \leq \sigma_2$.

---

61 This property is shared only by normal and stable laws, and is the main advantage of the use of stable laws for portfolio returns.
Therefore, a *scale parameter* can be chosen as a *risk measure* among stable distributions with the same $\alpha$ and $\beta$.

Since the $p$-th moments of $\alpha$-stable laws ($p < \alpha \leq 2$) are strictly increasing functions of a scale parameter, they can serve as risk measures as well.

Stable laws have a *domain of attraction*. Let $X_1, X_2, \ldots, X_n$ be independent copies of a random variable $X$ with distribution function $F$. Then $X$ is said to be in the domain of the attraction of a stable random variable $R$ (or its cdf $F$) if there exist $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$\frac{1}{a_n} \sum_{i=1}^{n} X_i - b_n \text{ converges in distribution to } R, \text{ as } n \to \infty,$$

$$\frac{1}{a_n} \sum_{i=1}^{n} X_i - b_n \rightarrow R.$$

Since the Pareto-stable distributions have infinite variances, one cannot estimate risk by variance and dependence by correlations. We shall introduce variance- and covariance-similar notions for stable laws. These notions are based on the multivariate assumptions of stable distributions.

**Definition 4**: A random vector $R$ of dimension $d$ is *stable* if for any $a > 0$ and $b > 0$ there exist $c > 0$ and a $d$-dimensional vector $D$ such that

$$aR_1 + bR_2 = cR + D$$

where $R_1$ and $R_2$ are independent copies of $R$.

If a random vector is stable with $\alpha > 1$, then it implies that all components of the vector are stable with the same index of stability and any linear combination (for example, portfolio returns) is again stable\(^{62}\).

The characteristic function of a $d$-dimensional vector is given by:

(a) if $\alpha \neq 1$,

---

\(^{62}\) We shall model the dependence structure of the vector of returns $(R_1, \ldots, R_d)$ of a portfolio by assuming that $(R_1, \ldots, R_d)$ is an $\alpha$-stable vector.
\[ \Phi_R(\theta) = \Phi_R(\theta_1, \theta_2, \ldots, \theta_d) = \]
\[ = \mathbb{E}\exp(i\theta^T R) = \exp\left\{ - \int_{S_d} |\theta^T s|^{\alpha} \left( 1 - i\text{sgn}(\theta^T s) \tan \frac{\pi \alpha}{2} \right) \Gamma(ds) + i\theta^T \mu \right\}, \]

(b) if \( \alpha = 1, \)
\[ \Phi_R(\theta) = \exp\left\{ - \int_{S_d} |\theta^T s| \left( 1 + i \frac{2}{\pi} \text{sgn}(\theta^T s) \ln |\theta^T s| \right) \Gamma(ds) + i\theta^T \mu \right\}, \]

where \( \Gamma \) is a bounded nonnegative measure on the unit sphere \( S_d, \) called a spectral measure, \( s \in S_d \) is the integrand unit vector, and \( \mu \) is the shift vector. If \( \alpha > 1, \) then \( \mu \) is the mean vector, \( \mu = \mathbb{E}R. \)

The scale parameter of a linear combination of the components of a stable vector \( R \) satisfies the relation:
\[ \sigma^\alpha(\theta^T R) = \sigma^\alpha(\theta_1 R_1 + \ldots + \theta_d R_d) = \int_{S_d} |\theta^T s|^{\alpha} \Gamma(ds). \]

Viewing \( R = (R_1, \ldots, R_d) \) as the vector of individual returns in a portfolio with weights \( \theta_1, \ldots, \theta_d, \) \( \sigma^\alpha(\theta^T R) \) will be the portfolio risk-measure.

As we defined above, \( \nu_\alpha = \sigma^\alpha \) is variation, a stable equivalent of the variance.

Similarly to traditional interpretation of a covariance as an indicator of dependence, one can use the covariation to estimate dependence between two \( s\alpha s \) distributions:
\[ [R_1; R_2]_\alpha = \frac{1}{\alpha} \left. \frac{\partial \sigma^\alpha(\theta_1 R_1 + \theta_2 R_2)}{\partial \theta_1} \right|_{\theta_1=0, \theta_2=1} = \int_{S_n} s^{\alpha-1} \Gamma(ds), \]

where \( (R_1, R_2) \) is a \( s\alpha s \) vector (\( 1 < \alpha \leq 2 \)) and \( x^{<k>} = \|x\|^k \text{sgn}(x) \) (signed power).

The matrix of covariations \([R_i; R_j]_\alpha, 1 \leq i \leq d, 1 \leq j \leq d,\) determines the dependence structure among the individual returns in the portfolio.
7.2. Modeling VAR with Stable Processes Exhibiting Time-varying Volatility

We shall employ stable Paretian distributions for approximating return distributions in VAR calculations. In this section, we consider conditional homoskedastic and conditional heteroskedastic return distributions based on normal and α-stable hypotheses. We shall compare VAR results obtained under different distributional modeling:

- ARMA(p,q) (Case 1),
- ARMA(p,q)-α-stable (Case 2),
- ARMA(p,q)-GARCH(r,s) (Case 3),
- ARMA(p,q)-stable-GARCH(r,s) (Case 4).

Larger numbers of cases indicate higher degrees of model generalization. We shall evaluate how generalization of methods improves VAR estimates. Case 1 is included as a benchmark model. Case 2 captures heavy tails in return distributions and reduces to Case 1 when the α parameter is 2. Cases 1 and 2 model the portfolio returns by i.i.d. innovation processes. The next step is to add a time-varying volatility to the model of asset returns. Case 3 describes time-varying volatility. Case 4 incorporates both thick tails and volatility changes.

7.2.1. Conditional Homoskedastic Distribution

Under the conditional distribution we imply the distribution, conditioned on the information available until the current moment t. In order to condition the process mean on past observations, researchers apply the autoregressive moving average (ARMA) models.

Suppose $R_1, R_2, ..., R_t, ...$ are observed returns on an individual asset. If the series show temporal dependence, then the ARMA(p,q) models are often used to describe the behavior of the data:

$$ R_t = \mu + \sum_{i=1}^{p} a_i R_{t-i} + \epsilon_t + \sum_{j=1}^{q} b_j \epsilon_{t-j}, \quad (7.1) $$
where \( \{\varepsilon_i\} \) is an innovation process.

We shall apply the ARMA model (7.1) under two assumptions for the innovation process:

- \( \{\varepsilon_i\} \) is a white noise process: \( \varepsilon_i \sim \text{iid N}(0, \sigma^2) \), (Case 1), and
- \( \{\varepsilon_i\} \) is iid \( \alpha \)-stable (Case 2).

In ARMA models the conditional variances are conditionally homoskedastic because they are independent of the previous observations. However, in practice the financial series are characterized by volatility clusters: one can distinguish periods of high volatility and tranquil periods.

### 7.2.2. Conditional Heteroskedastic Distribution

Time-varying volatility (conditional heteroskedasticity) can be captured by ARCH and GARCH models\(^{63}\). The ARMA(p,q)-GARCH(r,s) models assume that the returns follow the process (7.1) and the innovation process follows:

\[
\varepsilon_i = \sigma_i u_t, \quad (7.2)
\]

where \( u_t \sim \text{N}(0,1) \) and

\[
\sigma_i^2 = \alpha + \sum_{i=1}^{s} \beta_i \varepsilon_{t-i}^2 + \sum_{j=1}^{r} \gamma_j \sigma_{t-j}^2, \quad \sigma_i > 0, \quad (7.3)
\]

Panorska, Mittnik and Rachev (1995) and Mittnik, Rachev, and Paolella (1996) suggested stable GARCH models with the return process (7.1) and the innovation process

\[
\varepsilon_i = \sigma_i u_t, \quad (7.2)
\]

where \( u_t \) is \( \alpha \)-stable with \( \alpha > 1 \) and

\[
\sigma_i = \alpha + \sum_{i=1}^{s} |\beta_i| \varepsilon_{t-i} + \sum_{j=1}^{r} \gamma_j \sigma_{t-j}, \quad \sigma_i > 0. \quad (7.4)
\]

In the further empirical analysis we shall call the innovation process (7.2)-(7.3) as Case 3 and the model (7.2)-(7.4) as Case 4.

8. Modeling VAR with Stable Processes Exhibiting Short- and Long-range Dependence (Fractional-stable GARCH)

The fractional-stable GARCH model has not been analyzed in the VAR literature. The model captures all observed phenomena in financial data: heavy tails, time-varying volatility, and short- and long-range dependence.

We shall apply fractional stable processes exhibiting short- and long-range dependence for modeling the return data. Assume that the drift of returns is explained by the ARMA model (7.1) and the innovation process follows (7.2). We consider two processes for the random "driver" $u_t$ in (7.2):

- a fractional Gaussian noise (Case 5) and
- a fractional stable noise (Case 6).

Below we give formal definitions of the following notions needed to define a fractional-stable GARCH process as a model for financial returns:

- self-similarity,
- stationary increments,
- $H$-sssi processes,
- a fractional Brownian motion,
- a fractional Gaussian noise,
- short- and long-range dependence,
- an $\alpha$-stable Lévy motion,
- a fractional stable noise.

**Definition 1:** The real-valued process $\{R(t), t \in T\}$ is self-similar with index $H > 0$ ($H$-ss) if for all $c > 0$, the finite-dimensional distributions of $\{R(ct), t \in T\}$ are the same as
the finite-dimensional distributions of \{c^HR(t), t \in T\}, where T is either \(\mathbb{R}, \mathbb{R}_+ = \{t: t \geq 0\}\) or \(\{t: t > 0\}\).

**Self-similar processes** are processes, which remain the distributional form under scaling of time and space.

**Definition 2:** A real-valued process \(\{R(t), t \in T\}\) has *stationary increments* if

\[\{R(t + h) - R(h), t \in T\} \overset{d}{=} \{R(t) - R(0), t \in T\}, \text{ for all } h \in T.\]

**Definition 3:** The process \(\{R(t), t \in T\}\) is *H-sssi* if it is *self-similar with index H* and has stationary increments\(^{64}\).

For given \(0 < \alpha < 2\) and \(H\), there exist many \(\alpha\)-stable H-sssi processes. If \(\alpha = 2\) (the Gaussian case), there is a unique H-sssi process, called fractional Brownian motion.

**Definition 4:** A Gaussian H-sssi process, \(0 < H \leq 1\), is called *fractional Brownian motion (FBM)* and is denoted \(\{B_H(t), t \in \mathbb{R}\}\)\(^{65}\).

Since fractional Brownian motion \(\{B_H(t), t \in \mathbb{R}\}\) has stationary increments, its increments

\[F_j = B_H(j+1) - B_H(j), j = \ldots, -1, 0, 1, \ldots,
\]

form a stationary sequence.

**Definition 5:** The sequence \(\{F_j, j \in \mathbb{Z}\}\) is called *fractional Gaussian noise (FGN)*.

**Definition 6:** In general, given a process \(\{R_t, t \in \mathbb{R}\}\) with stationary increments and a noise process \(Y_t = R_t - R_{t-1}, R_0 = 0\), we say that \(Y_t\) is a fractional noise if \(R_t\) is a H-sssi process.

Let \(\gamma(j)\) be the autocovariance function of FGN, \(\gamma(j) = \mathbb{E}(F_0F_j), j \in \mathbb{Z}\).

**Definition 7:** The FGN exhibits:

- *short-range dependence (short memory)* if \(\sum_{j=-\infty}^{\infty} \gamma(j) < \infty\)

- *long-range dependence (long memory)* if \(\sum_{j=-\infty}^{\infty} \gamma(j) = +\infty\).

**Definition 8:** A process \(\{R(t), t \in \mathbb{R}\}\) with stationary strictly \(\alpha\)-stable independent

\(^{64}\) Example: Brownian motion is 1/2-sssi.

\(^{65}\) Example: When \(H=1/2\), fractional Brownian motion \(\{B_H(t), t \in \mathbb{R}\}\) is Brownian motion.
increments, $0 < \alpha < 2$, is called an \textit{\alpha-stable Lévy motion}\textsuperscript{66}.

The $\alpha$-stable Lévy motion is H-ssi with $H = \frac{1}{\alpha} \in \left( \frac{1}{2}, \infty \right)$.

\textbf{Definition 9}: A \textit{fractional stable noise} is the stationary sequence of the increments of the H-ssi $\alpha$-stable processes with a given $0 < \alpha < 2$.

9. Gaussian VAR Methods Based on Implied Volatility
10. VAR for Portfolios Containing Derivatives and the Use of Implied Distributions and Implied Tree
11. Comparison of the new methodologies with the existing methods
12. Conclusions

\textsuperscript{66} Example: The 2-stable Levy motion is Brownian motion.
References


8. Basle Committee on Banking Supervision, 1996, "Amendment to the Capital Accord to Incorporate Market Risks".


77. Linsmeier, Thomas and Neil Pearson, 1996, “Risk Measurement: An Introduction to Value at Risk”, University of Illinois at Urbana-Champaign, Department of Accountancy and Department of Finance.


Financial Engineering and Japanese Markets.


125. Venkataraman, Subu, 1997, “Value at Risk for a Mixture of Normal Distributions:

126. Wilson, Thomas, 1994, "Plugging the GAP”, *Risk* 7,10 (October), 74-80.

127. Zangari, Peter, 1996a, "A VAR Methodology for Portfolios that Include Option", *RiskMetrics Monitor*, (First Quarter 1996), 4-12.
