Lecture 7 - June 10

VAR - Market Risk (cont’d)
Portfolio Risk - Markowitz again!

• Portfolio VAR using information on components and correlations
• Will discuss incremental VAR, marginal VAR, component VAR
• All essential in practice to make risk management “actionable”
Portfolio VAR - components

• Define RoR \((t, t+1)\)

\[
R_{p,t+1} = \sum_{i=1}^{N} w_i R_{i,t+1}
\]

N-number of assets; \(R_{i,t+1}\) - ror on asset \(i\); \(w_i\)-weights

Portfolio return:

\[
R_p = \sum_{i=1}^{N} w_i R_i \equiv [w_1, w_2, w_3...w_N] \cdot w^T R
\]

Exp. Return:

\[
E(R_p) = \mu_p = \sum_{i=1}^{N} w_i \mu_i
\]
Portfolio VAR (cont’d)

- Exp variance  \( V(R_p) = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij} = \sum_{i=1}^{n} w_i^2 \sigma_i^2 + 2 \sum_{i=1}^{n} \sum_{j<i}^{n} w_i w_j \sigma_{ij} \)

\[ \sigma_p^2 = w^t \Sigma w \]

Need to translate portfolio variance into a VAR measure

If all security returns are normally distributed, then any linear combination of security returns will be normally distributed, so we can translate confidence level \( c \) into a normal deviate \( \alpha \)

\[ VAR_p = \alpha \sigma_p W = \alpha \sqrt{x^t \Sigma x} \]

Remember \( \sigma_{ij} = \sigma_i \sigma_j \rho_{ij} \)

\( \rho_{ij} \in [-1,1] \)
Portfolio VAR (cont’d)

• When we solved for the efficient frontier we applied the same equations- minimizing Portfolio variance for each return and finding the optimal weightings (Markowitz)

• One way of minimizing portfolio risk is to invest in assets which have low correlations to one another. Naturally, the larger the number of assets, the lower the overall portfolio risk

• Assume that a portfolio contains N assets, each with the same risk, all correlations the same and identical weights

• Then portfolio risk becomes

\[
\sigma_p = \sigma \left[ \frac{1}{N} + (1 - \frac{1}{N}) \rho \right]^{\frac{1}{2}}
\]

\[
\lim_{n \to \infty} (\sigma_p) = \sigma \sqrt{\rho}
\]
Marginal VAR

• Marginal contribution to risk by adding one additional security \( i \) with \( (\mu_i, \sigma_i, w_i) \) can be obtained:

\[
\frac{\partial \sigma^2_p}{\partial w_i} = 2w_i \sigma^2_i + 2 \sum_{i=1, j \neq i}^{N} w_j \sigma_{i j} = 2 \text{cov}(\mu_i, w_i \mu_i + \sum_{j \neq i}^{N} w_j \mu_j) = 2 \text{cov}(\mu_i, \mu_p)
\]

But we need the volatility’s contribution (not variance’s)

\[
\frac{\partial \sigma_p}{\partial w_i} = \frac{\text{cov}(\mu_i, \mu_p)}{\sigma_p}
\]

And further converting into VAR (therefore obtaining marginal VAR)

\[
\Delta VAR_i = \frac{\partial VAR}{\partial w_i W} = \alpha \frac{\partial \sigma_p}{\partial w_i} = \alpha \frac{\text{cov}(\mu_i, \mu_p)}{\sigma_p}
\]
Marginal VAR (cont’d)

• Remember “Beta” from CAPM?

\[ \beta_i = \frac{\text{cov}(\mu_i, \mu_p)}{\sigma_p^2} = \frac{\sigma_{ip}}{\sigma_p} = \rho_{ip} \frac{\sigma_i}{\sigma_p} = \rho_{ip} \frac{\sigma_i}{\sigma_p} \]

Beta measures the “systematic” risk of security I within portfolio p; using matrix notation,

\[ \beta = \frac{\sum W}{(W'\sum W)} \]

Therefore the relationship btw. Delta VAR and Beta becomes:

\[ \Delta VAR = \alpha (\beta_i \times \sigma_p) = \frac{VAR}{W} \times \beta_i \]
Incremental VAR

• Estimates the total impact of a proposed deal (trade) upon the existing portfolio \( p \)
• Ideally, Incremental VAR = \( \text{VAR}_{p+a} - \text{VAR}_p \)
• If \( \text{VAR} < \text{VAR} \) (before the deal), then the deal has a risk reducing effect (is a partial hedge) & if \( \text{VAR} \) (after the deal is completed) = 0 the deal is a total hedge
• Difficult to do because it requires a full reevaluation of the entire portfolio
• So we’ll expand VAR into a Taylor series and approximate the incremental VAR to the first derivative term:
• \( \text{VAR}_{p+a} = \text{VAR}_p + (\Delta \text{VAR})' \times a + \ldots \), so incr. VAR is prop. to \( \Delta \text{VAR} \)
When are we concerned about incremental VAR?

- In general when a new trade involves a set of new exposures on the risk factors (already within portfolio)

Flow chart for estimating incremental VAR

1. Initial portfolio \( p \)
2. Est future cov. matrix
3. Valuation
4. Portf. \( p \)
5. Portfolio with additional trade \( a \)
6. Portf. \( P+a \)
7. VAR(\( p \))
8. Del VAR
9. Incremental VAR
Component VAR

• Which exposure contributes how much to my overall portfolio VAR? (key question to make risk management “actionable”)

• If we multiply the marginal VAR by the dollar current position in asset or risk factor i, then: $Component_{VAR} = (\Delta VAR_i) \times w_i \Delta = VAR \beta_i w_i$

Thus Component VAR indicates how portfolio VAR would change (approximately) if the component were deleted from the portfolio.

Now we need to prove that the component VARs add up to the portfolio VAR: $CVAR_1 + CVAR_2 + ... + CVAR_N = VAR(N \sum_{i=1}^{N} w_i \beta_i) = VAR_i$

$CVAR_i = VARw_i \beta_i = (\alpha \sigma_i w_i \Delta) \rho_i = VAR_i \rho_i$

And upon normalizing by the total portfolio VAR we obtain the percent contribution to VAR of component i $\frac{CVAR_i}{VAR} = w_i \beta_i$
Principal Component Decomposition as an Eigenvalue Problem

- Consider a set of $N$ variables $\mu_1, \ldots, \mu_N$ with covariance matrix $\Sigma$
- We wish to reduce the dimensions of $\Sigma$ without too much loss of content, by approximating it with another matrix $\Sigma^*$
- Goal is to provide a good approximation of the portf. variance $z = w^T \mu$, using $V^*(z) = w^T \Sigma^* w$
- The process consists of replacing the original variables $\mu$ by another set $y$ suitably selected
- First principal comp: is the linear combination $y_1 = \beta_{11} \mu_1 + \ldots + \beta_{N1} \mu_N = \beta_1^T \mu$ such that its variance is maximized subject to a normalization constraint on the norm of the factor exposure vector $\beta_1^T \beta_1 = 1$
Principal Component Decomposition (cont’d)

• A constrained optimization of this variance \( \sigma^2(y_1) = \beta_1' \Sigma \beta_1 \) shows that the vector \( \beta_1 \) must satisfy \( \Sigma \beta_1 = \lambda_1 \beta_1 \)
• Here \( \sigma^2(y_1) = \lambda_1 \) is the largest eigenvalue of matrix \( \Sigma \) and \( \beta_1 \) its associated eigenvector
• The second principal component is the one that has greatest variance subject to the same normalization constraint \( \beta_2' \beta_1 = 1 \) and to the fact that it must be orthogonal to the first \( \beta_2' \beta_1 = 0 \)
• And so on for all the others
• So we managed to replace the original set of variables \( \mu \) by another set of \( y \) orthogonal factors with the same dimension but sorted by decreasing importance
Principal Component Decomposition (cont’d)

• This leads to the singular value decomposition

\[ \Sigma = PDP' = [\beta_1 \ldots \beta_N] \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_N \end{bmatrix} [\beta'_1 \ldots \beta'_N] \]

PP’=I (orthogonal matrix); D diagonal matrix of \( \lambda_j \)

To each \( y_j \) we associated a value for its variance \( \lambda_j \) sorted in order of decreasing importance. The eigenvalues are an indicator on the “true” dimensionality of the \( \Sigma \) matrix

If the eigenvalues are small then the respective risk factors are negligible

But what if they are negative? (problem of non positive definition of covariance matrix - very real, unfortunately)
Dealing with non positive definite covariance matrixes - a practical point of view

- We decide to keep only the first K components beyond which their variances are deemed as “unimportant”
- We approximate thus the previous linear combination by
  \[ \mu_i \approx \beta_{i1} y_1 + \ldots + \beta_{ik} y_{ik} \]

\[
\Sigma^* = \begin{bmatrix} \lambda_1 \ldots 0 \\ \vdots \\ 0 \ldots \lambda_k \end{bmatrix} \begin{bmatrix} \beta_{1}' \\ \vdots \\ \beta_{k}' \end{bmatrix} = \beta_1 \beta_1' \lambda_1 + \ldots + \beta_k \beta_k' \lambda_k
\]

Observe that \( \Sigma^* \) is not invertible since \( k < N \)

In practice we map our portfolios into exposures to principal components until their effects become negligible, but weight the exposures by PC
Forecasting variances and correlations (determining the appropriate $\Sigma$ matrix)

- Working with the appropriate correlation matrix is key to successfully assessing risk
- The problem is far from trivial:
  - variances and correlations vary over time
  - historical values are often not representative for future expectations
  - interpolations are very tricky, unlike yield curves there’s nothing calling for smooth volatility curves
Forecasting volatilities and correlations - a parametric approach

• The crudest yet often used method is of moving averages \( \sigma_t^2 = \frac{1}{M} \sum_{i=1}^{M} \mu_{t-i}^2 \)

Notice that we focus on raw returns as opposed to returns around the mean

Drawbacks: ignores the dynamic ordering of observations - recent observations are equally weighted as historical observations of less relevance (volatility tends to behave in “buckets”)

One high return skewes the estimates unnecessarily
GARCH Estimation

• GARCH - generalized autoregressive heteroskedastic model (Engle-1982 & Bollerslev - 1986) (heteroskedastic means that variances are changing - as you know from your regression courses)

• Assumes that the return variances follow predictable processes

• The conditional variance depends on the latest innovation and on the previous conditional variance

• Define $h_t$ as the conditional variance using information up to $t-1$

• Define $\mu_{t-1}$ the previous day’s return

• GARCH (1,1): $h_t = \alpha_0 + \alpha_1 r^2_{t-1} + \beta h_{t-1}$

The average, unconditional variance is found by setting $E(r^2_{t-1}) = h_t = h_{t-1} = h$

$$h = \frac{\alpha_0}{1 - \alpha_1 - \beta}$$

To be stationary $\alpha_1 + \beta < 1$
Problems with GARCH

• In practice GARCH exhibits (in one or another variant) good results, but:
• It is heavily non-linear (i.e. the parameters must be estimated by maximization of a likelihood function involving numerical optimization - see your courses on numerical methods and operations research)
• Typically it is assumed that the scaled residuals $\varepsilon_t = \frac{\mu_t}{\sqrt{h_t}}$

Are normally distributed and independent; if T observations, their joint density is the product of the densities for each t

Objective function (f is the normal density function):

$$\max F(\alpha_0, \alpha_1, \beta | \mu) = \sum_{t=1}^{T} \ln f(\mu_t | h_t) = \sum_{t=1}^{T} \left( \ln \frac{1}{\sqrt{2\pi h_t}} - \frac{\mu_t^2}{2h_t} \right)$$
The Risk Metrics Approach

• Risk Metrics originally developed by JP Morgan (1995), now supported by Reuters
• Pragmatic approach using exponentially weighted moving averages
• For goods and bads see attached articles (“How safe is RiskMetrics? - Colin Lawrence & Gary Robinson) and the “response” “A transparent tool” - Jacques Longestraey & Peter Zangari), also the JP Morgan Risk Metrics technical document available on the internet
• Forecast for time t is a weighted average of the previous forecast, using weight $\lambda$ (the decay factor), and of the latest squared innovation, using weight $(1-\lambda)$:

$$h_t = \lambda h_{t-1} + (1-\lambda)\mu^2_{t-1}$$

$\lambda < 1$

$$h_t = (1-\lambda)(\mu^2_{t-1} + \lambda\mu^2_{t-2} + \lambda^2\mu^2_{t-3} + ...)$$
Market VAR Conundrum: Local vs. Full Valuation

- Local valuation methods measure risk by valuing the portfolio once, at the initial position, and using local derivatives to infer possible movements (remember last lecture - the Taylor expansion of VAR?)
  - Delta normal approach stops at the first derivatives
  - “Greeks” use also the second order approximations but neglect a number of phenomena which can carry importance at tail end events
Full valuation methods measure risk by fully repricing the portfolio over a range of scenarios
- Historical Simulations
- Monte Carlo Simulations
On the “Delta - Gamma” Approximations

- The details of performing full valuation risk estimation were discussed last session.
- So let’s discuss the Taylor expansion with the second order terms included:

\[
dV = \frac{\partial V}{\partial S} dS + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} dS^2 + \frac{\partial V}{\partial t} dt + \ldots = \Delta dS + \frac{1}{2} \Gamma dS^2 + \Theta dt + \ldots
\]

Taking variances on both sides, we obtain:

\[
\sigma^2(dV) = \Delta^2 \sigma^2(dS) + \left(\frac{1}{2} \Gamma^2\right) \sigma^2(dS^2) + 2(\Delta \Gamma) \text{cov}(dS, dS^2)
\]

Now, if dS is normally distributed, all odd moments vanish and therefore so does the last term; likewise \( V(dS^2) = 2V(dS)^2 \).

Assuming that dS and dS^2 are jointly normally distributed, then dV is normally distributed and VAR can be calculated parametrically

\[
VAR = \alpha \sqrt{\left(\Delta S \sigma\right)^2 + \left(1/2\right)\left(\Gamma S^2 \sigma^2\right)^2}
\]

\( \alpha \) and \( \sigma \) are parameters.
On the “Delta - Gamma” Approximations (cont’d)

• Naturally, if dS is a normal, dS^2 can’t also be normal (typically χ^2), therefore last page was only a rough approximation

• One can further account for the skewness coefficient \( \xi \) and the corrected VAR is derived by replacing \( \alpha \) by \( \alpha’ = \alpha -1/6(\alpha^2-1) \xi \) (derived by using the Cornish Fisher expansion)

• Obviously there is no correction under a normal distribution, where \( \xi = 0 \)

• A numerically more elaborate method uses Monte Carlo simulations of the random variable S (risk factors) and then uses the Taylor expansion to second order
When to use which method?

- For large portfolios where optionality is not a dominant factor, the delta normal method is fast, efficient and pretty accurate.
- For portfolios exposed to a few sources of risk and with substantial optionality, the Delta-Gamma-Theta provides significant improvement in precision at little incremental computational cost.
- For portfolios with substantial options components (like mortgages) or longer horizons, best perform a full valuation approach.
- Most sizeable international banks employ full valuation methods for their enterprise risk management purposes as well as market risk capital allocation.
- At HVB Group we also perform full valuation within a Monte Carlo simulation framework.
Full Valuation - Historical Simulations

- Select a sample of actual daily risk factor changes over a given period of time (Germany- 250 days)
- Apply those changes to the current value of the risk factors and revalue the portfolio as many times as the number of days in the historical sample
- Construct the histogram of portfolio values and identify the VAR that isolates the first percentile of the distribution in the left hand tail (assuming we use a 99% confidence level, if not $\alpha$)
- Is very attractive because:
  - involves scenarios which have happened such that no one can argue against the chosen scenarios
  - Is completely non parametric
  - To pros and cons a little later
Full Valuation - Monte Carlo Simulations

• Consists of repeatedly simulating the random processes that govern market prices and rates (therefore our risk factors)
• Each simulation (scenario) generates a possible value of the portfolio at the target horizon (e.g. 1 day or 10 days)
• If we generate enough scenarios the simulated distribution of the portfolios values will converge towards the true, although unknown, distribution (remember second lecture on applying the central limit theorem and Chebyshev inequality to prove convergence of MC method?)
• The VAR can be read off the distribution like in the case of historical simulations
Monte Carlo Simulations (cont’d)

- So first one needs to specify all relevant risk factors which impact the portfolio’s value (vertex points on curves, volatilities, correlations (in case of quantos, etc.), FX spot rates, asset prices, etc.)
- Hull (2000) offers a detailed analysis of the stochastic processes which model the behavior of various financial assets (like geometric brownian motion, etc.)
- Secondly, we need to construct price paths
  - when several correlated risk factors are involved we need to simulate multivariate distributions
  - Typically the risk factors (or assets) are correlated; Assuming a series of stocks $S_i$, each following Wiener processes $(\mu_i, \sigma_i)$ with returns correlated by $\rho_{ij}$, meaning that $E[W_t^i, W_t^j] = \rho_{ij} t$, then the multivariate process becomes
  \[
  S'_t = S'_0 \exp\left(\left(\mu_i - \frac{1}{2} \sigma^2_i\right)t + \sqrt{t} \sigma_i X_i\right),
  \]
  \[
  X \approx N(\sigma, \Sigma) | \Sigma_{ij} = E(XX^T) = \rho_{ij}
  \]
Monte Carlo Simulations (cont’d)

• Therefore we need to start by generating n independent random variables \((Y=(Y^1_t,\ldots,Y^n_t))\) with mean zero and variance 1, so that \(\text{E}[YY^T]=I\)

• Next, construct the variables \(X=AY\), where \(A\) is a matrix such that \(\Sigma=AA^T\)

• Then value the portfolio for each path (scenario)

• 99% VAR is then defined as the distance to mean of the first percentile of the distribution

• Naturally, like in the case of historical simulation you need to have a pricing function for every security, difficult task because securities tend to deviate from equilibrium pricing models (where they exist) for large instantaneous movements in the risk factors
Pros and Cons of the parametric (variance - covariance) method

**PROS**

- computationally efficient
- works well even when risk factors are not normally distributed (provided they are numerous and relatively independent)
- doesn’t require pricing models, only the greeks (typically easily provided by modern front office trading systems)
- Easy to do incremental, marginal VAR and factor de-composition

**CONS**

- Assumes normality of the returns
- Assumes that the risk factors follow a multivariate log normal distribution, therefore neglects fat tails
- Requires estimation of volatilities of risk factors and correlations of returns, tricky as we’ve seen
- Limits valuation to second order expansion (and even there neglecting correlations among the greeks)
- Can’t be used to conduct sensitivity analysis
Pros and Cons of historical simulations

**PROS**
- no need to make assumptions on the distributions of risk factors
- no need to forecast volatilities or correlations (implicitly captured by the “real” synchronous movements in daily prices)
- Fat tails are captured so long they’re present in the data series used for the simulations
- Aggregation across markets is straightforward
- Allows calculation of confidence intervals

**CONS**
- Dependence on data set and its idiosyncrasies. Extreme events either in or out.
- Can’t accommodate changes in the market structure, like the introduction of the Euro
- Short data set may not be reflective of future events (bias)
- Can’t be used for sensitivity analyses
- Not always computationally efficient when portfolio contains complex securities
Pros and Cons for Monte Carlo simulations

• PROS
  • Accommodates virtually any distribution of risk factors
  • Can be used to model any complex portfolio
  • Allows the calculation of confidence intervals for VAR
  • Allows for sensitivity analyses and stress testing

• CONS
  • Outliers aren’t incorporated into the distribution
  • Computationally intensive and expensive in hardware and software
Additional tools to capture market risk

• Back testing
• required by regulators
• discussed last time
• Stress testing
• Incorporation of extreme scenarios typically not captured by any other means
• Model weakness analyses - prediction power
Next Lecture - 17 June

• Credit Risk