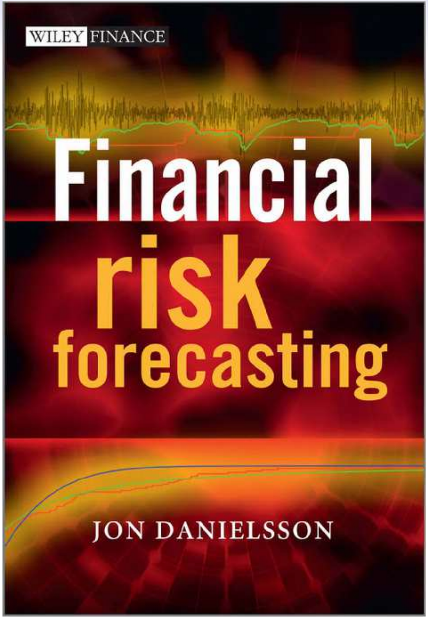
Financial Risk Forecasting
Chapter 3
Multivariate Volatility Models

Jon Danielsson ©2025
London School of Economics

To accompany
Financial Risk Forecasting
FinancialRiskForecasting.com

Published by Wiley 2011

Version 10.0, August 2025


WILEY FINANCE
Financial
risk
forecasting
JON DANIELSSON

Multivariate Volatility Models

Why Multivariate Volatility

- In practice, we hold portfolios, not single assets
- We need to estimate not just individual volatilities, but also correlations between assets
- These relationships are summarised in the conditional covariance matrix
- Multivariate volatility models are therefore essential for portfolio risk management, hedging, and scenario analysis
- However, modelling many assets jointly introduces significant complexity

What This Chapter Covers

- Multivariate volatility modelling approaches
- EWMA (Exponentially Weighted Moving Average)
- CCC (Constant Conditional Correlation)
- DCC (Dynamic Conditional Correlation)
- Approaches for large systems
- Model comparison and implementation

Notation new to this Chapter

Σ_t	Conditional covariance matrix
$y_{t,i}$	Return on asset i at time t
$y_t = (y_{t,1}, \ldots, y_{t,K})$	Vector of asset returns at time t
$y = (y_1, \ldots, y_T)'$	Matrix of returns across assets and time
R	Correlation matrix
D_t	Conditional variance forecast

R Estimation

- It is easy to implement EWMA directly in R
- No single package implements all models we will cover
- We mostly use Alexios Ghalanos's `rmgarch`

Other Reading

- Many surveys, for example
- Boudt C., A. Ghalanos, S. Payseur, and E. Zivot. "Multivariate GARCH models for large-scale applications: A survey". Handbook of Statistics. Elsevier
- For large (more than 25 assets) this paper has a proposal
- Engle, R. F., Pakel, C., Shephard, N., and Sheppard, K. (2019). "Fitting vast dimensional time-varying covariance models."

Learning outcomes

1. Understand the key issues in the forecasting of the conditional covariance matrix
2. Know the curse of dimensionality problem
3. Recognise the importance of positive definiteness of the conditional covariance matrix
4. Derive the multivariate EWMA model
5. Derive the CCC and DCC models and recognise the strengths and weaknesses of each
6. Implement the estimation of multivariate volatility models in R
7. Apply scalable techniques to high-dimensional volatility modelling problems

Multivariate Volatility Forecasting

From Univariate to Multivariate Volatility

- So far, we have looked at models for the volatility of a single asset
- In practice, portfolios consist of many assets, not just one
- We need to model how each asset varies and how they move together
- This leads us to the conditional covariance matrix of returns

Setup

- Consider the univariate volatility model:

$$y_t = \sigma_t \epsilon_t,$$

where y_t are returns, σ_t is conditional volatility and ϵ_t are random shocks

- If there are $K > 1$ assets under consideration, it is necessary to indicate which asset and parameters are being referred to, so the notation becomes more cluttered:

$$y_{t,i} = \sigma_{t,i} \epsilon_{t,i},$$

where the first subscript indicates the date and the second subscript the asset

Conditional Covariance Matrix Σ_t

- The conditional covariance between two assets *i* and *j* is indicated by:

$$\text{Cov}(y_{t,i}, y_{t,j}) \equiv \sigma_{t,ij}$$

- In the three-asset case (note that $\sigma_{t,ij} = \sigma_{t,ji}$):

$$\Sigma_t = \begin{pmatrix} \sigma_{t,11} & & \\ \sigma_{t,12} & \sigma_{t,22} & \\ \sigma_{t,13} & \sigma_{t,23} & \sigma_{t,33} \end{pmatrix}$$

Portfolio Variance

- Let w be a vector of portfolio weights
- The portfolio variance is

$$\sigma^2_{\text{portfolio}} = w'\Sigma w$$

- Estimating Σ_t is the main goal of multivariate volatility modelling

The Curse of Dimensionality

- Covariance matrix has K variances and $K(K-1)/2$ covariances
- Total parameters to estimate:

$$K + \frac{K(K-1)}{2}$$

- Grows quickly: 2 assets $\rightarrow$ 3 terms, 3 assets $\rightarrow$ 6 terms, 10 assets $\rightarrow$ 55 terms
- This complexity makes full estimation hard as K increases

Positive Semi-Definiteness

- For univariate volatility, we need to ensure that the variance is not negative ($\sigma^2 \geq 0$)
- And for a portfolio

$$\sigma^2_{\text{portfolio}} = w'\Sigma w \geq 0$$

- A covariance matrix must be positive semi-definite

$$w'\Sigma w \geq 0 \quad \forall\ w \in \mathbb{R}^K$$

- It is not easy to guarantee this in practice

What About a Full MV-GARCH Model?

- For one asset

$$\sigma^2_{t+1} = \omega + \alpha y^2_t + \beta \sigma^2_t$$

- For two

$$\sigma_{t+1,11} = \omega_1 + \alpha_1 y^2_{t,1} + \beta_1 \sigma_{t,11} + \alpha_2 y^2_{t,2} + \beta_2 \sigma_{t,22} + \delta_1 \sigma_{t,1,2} + \gamma_1 y_{t,1} y_{t,2}$$
$$\sigma_{t+1,22} = \omega_2 + \alpha_3 y^2_{t,1} + \beta_3 \sigma_{t,11} + \alpha_4 y^2_{t,2} + \beta_4 \sigma_{t,22} + \delta_2 \sigma_{t,1,2} + \gamma_2 y_{t,1} y_{t,2}$$
$$\sigma_{t+1,1,2} = \omega_3 + \alpha_5 y^2_{t,1} + \beta_5 \sigma_{t,11} + \alpha_6 y^2_{t,2} + \beta_6 \sigma_{t,22} + \delta_3 \sigma_{t,1,2} + \gamma_3 y_{t,1} y_{t,2}$$

- Or 21 parameters to estimate
- Almost impossible in practice

Numerical Issues

- Violation of covariance stationarity usually don't matter for univariate GARCH
- A univariate volatility forecast is still obtained even if $\alpha + \beta > 1$
- Multivariate models are less forgiving
- Stationarity and positive definiteness must be explicitly enforced

Numerical Issues (Cont.)

- A parameter set resulting in violation of covariance stationarity might also lead to unpleasant numerical problems
- Numerical algorithms need to address these problems, thus complicating the programming process considerably
- Problems with multiple local minima, flat surfaces and other pathologies discussed in the last chapter

Feasible Alternatives to Full MV-GARCH

- Use some simplification approaches
- Unfortunately they come with significant trade-offs
- So MV estimation is much harder and much less accurate than univariate estimation
  1. EWMA
  2. CCC
  3. DCC (perhaps most widely used for portfolios)
  4. Hierarchical models (perhaps most widely used for combining portfolios)

  Common models not covered here

  5. BEKK (not to be recommended except for very small problems, $K = 2$, perhaps $K = 3$)
  6. MV GARCH (practically impossible except maybe when $K = 2$)
  7. Composite likelihood

EWMA

EWMA

- The Exponentially Weighted Moving Average (EWMA) model is one of the simplest ways to forecast volatility
- It places more weight on recent observations and less on older ones
- In the multivariate case, EWMA provides a simple method to estimate a time-varying covariance matrix
- Despite its simplicity, it is widely used in risk management
- A practical starting point before moving to more flexible models like DCC

EWMA Model

- Univariate:

$$\hat{\sigma}_t^2 = \lambda \hat{\sigma}_{t-1}^2 + (1-\lambda) y_{t-1}^2$$

- A vector of returns is

$$\underset{K \times 1}{y_t} = [y_{t,1}, y_{t,2}, \ldots, y_{t,K}]'$$

- The multivariate EWMA is:

$$\hat{\Sigma}_t = \lambda \hat{\Sigma}_{t-1} + (1-\lambda) y_{t-1} y_{t-1}',$$

with an individual element given by:

$$\hat{\sigma}_{t,ij} = \lambda \hat{\sigma}_{t-1,ij} + (1-\lambda) y_{t-1,i} y_{t-1,j}$$

Properties

- The same weight, λ, is used for all assets
- It is pre-specified and not estimated
- Each element of the covariance matrix depends only on its own past values and the corresponding returns

Strengths of Multivariate EWMA

- Straightforward implementation, even for a large number of assets
- Covariance matrix is guaranteed to be positive semi-definite
- Computationally light and easy to update in real time
- Used in practice by many risk systems as a baseline model

Weaknesses of Multivariate EWMA

- Simple structure — assumes the same dynamics for all assets
- Single decay factor λ used for all series
- λ is usually fixed, not estimated from data
- Cannot adapt to changing correlations over time, limiting its accuracy in volatile market regimes

Constant Conditional Correlation Models

Constant Conditional Correlation (CCC) Model

- The CCC model (Constant Conditional Correlation) is a multivariate GARCH framework
- Assumes conditional correlations between asset returns remain *constant* over time
- It separates volatility and correlation modeling: each asset's volatility is modeled individually (e.g., with its own GARCH model)
- A single fixed correlation matrix captures their co-movement

Constant Conditional Correlation (CCC) Model (con't)

- By keeping the correlation matrix fixed
- The CCC model is much simpler to estimate (fewer parameters) and easier to implement than models with time-varying correlations
- A baseline for multivariate volatility analysis when correlations are believed to be stable
- Also provides a benchmark for evaluating more advanced models that allow *time-varying* correlations (like DCC)

Steps

- We usually start by removing the mean — de-mean returns
- Model volatilities and correlations in two steps, assuming independence between steps (known as the two-stage procedure)
  1. Correlation matrix
  2. Variances
- Model volatilities with GARCH or some standard method
- The correlation matrix is static

Definitions

- Let D_t be a diagonal matrix where each element is the *volatility* of each asset

$$D_{t,ii} = \sigma_{t,i}, \quad i = 1, \ldots, K$$
$$D_{t,ij} = 0, \quad i \neq j$$

- Let R be the *unconditional* correlation matrix
- We employ a two-step procedure
  1. Estimate D_t
  2. Estimate R
- Then combine them to get the conditional covariance matrix, Σ_t

Step 1. Conditional Variance

- Use univariate GARCH (or some method) to estimate the variance of each asset separately, perhaps

$$\sigma^2_{t,i} = \omega_i + \alpha_i y^2_{t-1,i} + \beta_i \sigma^2_{t-1,i}$$

- Note the parameters will be different for each asset
- D_t is then created by putting these into the diagonal elements

$$D_{t,ii} = \sigma_{t,i}, \quad i = 1, \ldots, K$$

Residuals

- GARCH model for each asset is

$$\sigma_{t,i}^2 = \omega_i + \alpha_i y_{t-1,i}^2 + \beta_i \sigma_{t-1,i}^2$$

- Or each asset is given by

$$y_{t,i} \sim \mathcal{N}(0, \sigma_{t,i}^2)$$

- The estimated residuals, $\hat{\epsilon}_{t,i}$

$$\hat{\epsilon}_{t,i} = \frac{y_{t,i}}{\hat{\sigma}_{t,i}} = \frac{y_{t,i}}{\sqrt{\hat{\omega}_i + \hat{\alpha}_i y_{t-1,i}^2 + \hat{\beta}_i \hat{\sigma}_{t-1,i}^2}}$$

Getting the Correlations, R

- The vector of residuals at time t is then

$$\underset{K\times 1}{\hat{\epsilon}_t} = \hat{D}_t^{-1} y_t$$

- And the matrix of residuals for all assets and times is

$$\underset{K\times T}{\hat{\epsilon}}$$

Residual Correlations

- We want the correlations of the residuals
- Since mean is zero and variance one, the $K \times K$ correlation matrix is the same as the covariance matrix

$$\hat{R} := \mathrm{Cov}(\hat{\epsilon}) = \frac{\hat{\epsilon}\hat{\epsilon}'}{T}$$

Constant Conditional Correlations (CCC)

- We now have the constant matrix $\hat{R}$ and the time varying matrix $\hat{D}_t$
- Combine these two to create the conditional variance matrix

$$\hat{\Sigma}_t = \hat{D}_t \hat{R} \hat{D}_t$$

Pros

1. Guarantees the positive definiteness of $\hat{\Sigma}_t$ if $\hat{R}$ is positive definite, as it is by definition
2. Simple model, easy to implement
3. Since matrix $\hat{D}_t$ has only diagonal elements, each univariate GARCH model can be estimated independently, simplifying computation

Problem 1: Constant Correlation Assumption

- The CCC model assumes that conditional correlations between asset returns are constant over time
- In reality, correlations are time-varying, especially during market stress or structural changes
- Consequently, it fails to capture shifts in systemic risk and market contagion
- For example, asset correlations often spike during financial crises and market stress, which CCC cannot accommodate

Problem 2: Misspecification Risk

- Even with correctly specified univariate GARCH processes, a fixed correlation matrix may result in misspecification of the joint distribution
- Consequences include
  - Inaccurate modelling of joint return dynamics
  - Poor estimation of portfolio volatility and Value-at-Risk (VaR) (see next chapters)
- Because the constant correlation assumption ignores evolving relationships among assets

Problem 3: Poor Fit for Heterogeneous Assets

- CCC performs poorly when applied to diverse assets (e.g., equities, bonds, commodities)
- Different asset classes exhibit different correlation patterns and dynamics
- The limitation is that CCC imposes the same correlation structure across all periods, regardless of asset behaviour
- Result: Forecasts become inefficient compared to more flexible models

CCC Summary

- Strength: Simplicity, tractability, positive definiteness
- Weakness: Cannot capture time-varying correlation structure
- Practical Use: Benchmark model; good for stress-testing assumptions
- Next: DCC models relax the constant correlation assumption

Dynamic Conditional Correlation Models

CCC to DCC

- The CCC model simplifies analysis by assuming that correlations are constant over time
- However, real-world data shows that asset correlations often change—especially during market stress or economic shifts
- A model with *time-varying correlations* is needed to better capture co-movements in financial markets
- This leads us to the *Dynamic Conditional Correlation (DCC)* model

CCC vs DCC: Summary

Feature	CCC Model	DCC Model
Volatility modelling	Univariate GARCH	Univariate GARCH
Correlation structure	Constant	Time-varying
Estimation complexity	Low	Moderate
Number of parameters	Fixed	Fixed (shared)
Asset-specific dynamics	No	No
Time adaptability	✗	✓
Common use	Benchmark	Portfolio modelling

Motivation for Dynamic Correlations

- In turbulent periods, correlations between assets can increase sharply — CCC cannot capture this
- DCC extends CCC by allowing the correlation matrix to evolve over time, while still using univariate GARCH models for volatilities
- The goal is to keep the model flexible enough to adapt to changing market conditions, yet simple enough to estimate efficiently
- DCC is widely used in finance for modelling portfolios, risk and contagion

Dynamic Conditional Correlations (DCC)

- DCC model is an extension of the CCC
- Suppose we let the conditional correlation matrix R_t be *time-dependent*
- Note we will also use the unconditional correlation matrix, R, from the last section
- Be alert to R and R_t

The obvious way

- The obvious way to model R_t is to do it like GARCH model
- To let correlations vary over time, we could use past shocks and past correlations to update our estimate of the current correlation matrix like

$$R_t = a + b\epsilon_{t-1}\epsilon'_{t-1} + cR_{t-1}$$

- That will not work (see next slide)

Correlation Matrix Requirements

- Positive definiteness (PD)

$$\sigma^2_{\text{portfolio}} = w'\Sigma w \geq 0$$

- D_t is positive definite by construction since all elements in D_t are positive or zero
- All elements of R_t need to be ≤ 1 and ≥ -1
- The diagonal elements are one
- And equation

$$R_t = a + b\epsilon_{t-1}\epsilon'_{t-1} + cR_{t-1}$$

- Does not ensure that

We Need More Steps

- Break R_t into two parts:

  Q_t A positive definite matrix that captures the raw time-varying dynamics of correlations

  Z_t A diagonal matrix that rescales Q_t — this ensures that the final matrix has valid correlations between -1 and 1

- We then combine these as:

$$R_t = Z_t Q_t Z_t$$

The Diagonal Elements of a Correlation Matrix Must be 1

- We show after this slide how to obtain Q_t
- Each diagonal entry of Z_t is the inverse square root of the corresponding entry in Q_t:

$$Z_t = \begin{pmatrix} 1/\sqrt{q_{t,11}} & 0 & \cdots & 0 \\ 0 & 1/\sqrt{q_{t,22}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1/\sqrt{q_{t,KK}} \end{pmatrix}$$

Modelling the Correlation Dynamics

- To allow correlations to change over time, we update Q_t using a GARCH-like rule
- First, we initialise Q_t on day $t = 1$ with the constant correlation matrix estimated from CCC:

$$Q_1 = \hat{R}$$

- Then, for each $t > 1$, we use the following update:

$$Q_t = (1 - \zeta - \xi)\,\hat{R} + \zeta\,\epsilon_{t-1}\epsilon'_{t-1} + \xi\,Q_{t-1}$$

- This formula combines:
  - The long-run average level of correlation ($\hat{R}$)
  - Recent shock information ($\epsilon_{t-1}\epsilon'_{t-1}$)
  - The previous estimate (Q_{t-1})
- Just like GARCH does for variance, this structure lets Q_t respond to recent surprises and gradually adapt over time

What the Parameters Mean

- The parameters ζ and ξ control how Q_t evolves:
  - ζ determines how much recent shocks affect Q_t
  - ξ controls how much weight is placed on past correlations versus new shocks
- The term $(1-\zeta-\xi)$ ensures that the matrix doesn't drift away from the long-run average
- To make sure Q_t remains valid (positive definite), we need:

$$\zeta, \xi > 0 \quad \text{and} \quad \zeta + \xi < 1$$

DCC Model Structure

- The DCC model is estimated in two stages:

Stage 1. Volatility estimation (perhaps univariate GARCH)

$$y_{t,i} = \sigma_{t,i}\epsilon_{t,i}, \quad \sigma^2_{t,i} = \omega_i + \alpha_i y^2_{t-1,i} + \beta_i \sigma^2_{t-1,i}$$

$\rightarrow$ gives $D_t = \text{diag}(\sigma_{t,1}, \ldots, \sigma_{t,K})$

Stage 2. Correlation estimation (DCC recursion)

$$Q_t = (1 - \zeta - \xi)\hat{R} + \zeta \epsilon_{t-1}\epsilon'_{t-1} + \xi Q_{t-1}$$

$$R_t = Z_t Q_t Z_t$$

- Combine to get full conditional covariance matrix:

$$\Sigma_t = D_t R_t D_t$$

Strengths of the DCC Model

- One of the biggest advantages of the DCC model is its scalability
- It allows us to estimate large conditional covariance matrices with relative ease
- Why this is useful:
    - In a portfolio with many assets, full multivariate GARCH models become too complex and parameter-heavy
    - DCC retains univariate GARCH modelling for volatilities while jointly modelling time-varying correlations, and only models the correlations jointly
- As a result, the DCC model is popular in empirical finance, where large datasets are common but overfitting must be avoided

Limitations of the DCC Model

- A key drawback of the standard DCC model is that it uses the same parameters ζ and ξ for all asset pairs
- This means:
  - All assets are assumed to respond to shocks and past correlations in the same way
  - The model imposes the same correlation dynamics across the board
- While this keeps estimation simple, it can be unrealistic in practice — different asset classes often behave differently
- More advanced models have been proposed to address these issues

Large Problems — Hierarchical Approach

The Limits of DCC in Large Systems

- The DCC model is efficient for moderate numbers of assets — but it doesn't scale well to very large systems
- A medium-sized financial institution might track hundreds of thousands of risk factors — across geographies, asset classes and business lines
- Estimating a full covariance matrix across all of these using DCC becomes infeasible
- Why?
  - Computational burden grows quadratically with the number of assets
  - i.e. covariance matrix has $K(K+1)/2$ parameters, growing as $\mathcal{O}(K^2)$
  - A single correlation structure across all assets is too restrictive
- We need a more structured approach to cope with this complexity

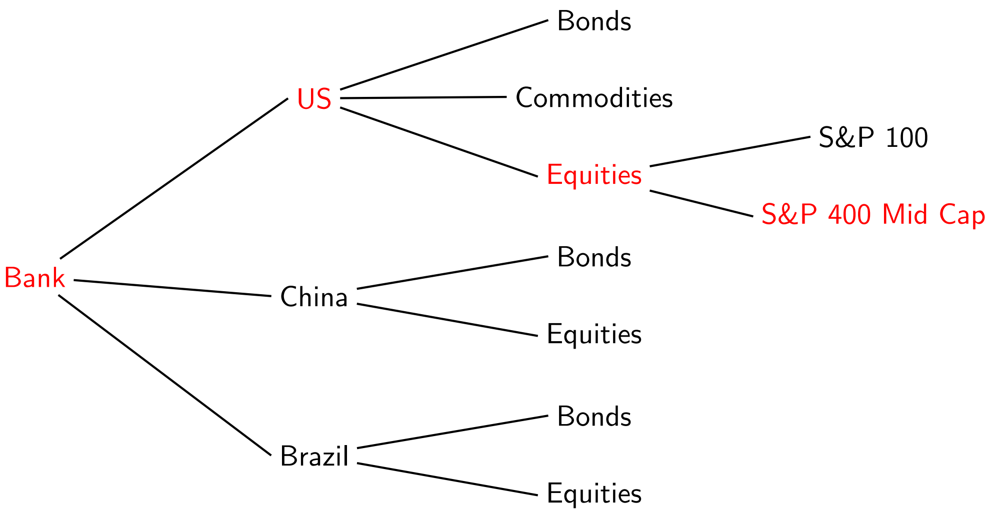
A Hierarchical Solution

- Instead of treating all assets as a single block, we can break them into meaningful groups
- Example groups: by region, by asset class, or by business unit
- Within each group, we can estimate volatility and correlation using a model like DCC or even simpler alternatives
- Then, we combine group-level outputs to construct the institution-wide covariance matrix
- This approach:
  - Reduces dimensionality
  - Captures local structure more accurately
  - Respects organisational boundaries

Example: Hierarchical Structure

- A large institution can be broken into logical subgroups for tractable modelling
- This tree shows an example hierarchy based on geography and asset class
- Each node can be modelled independently, and risks aggregated hierarchically
- We aim to estimate volatility and correlation within and across these groupings

Example: Tree Representation



Simplifying the Problem

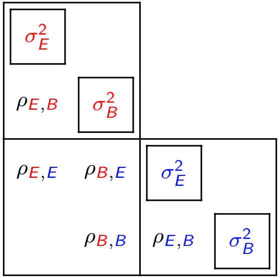
- Suppose you are willing to assume a single correlation between all Brazilian assets and all US assets
- And you are willing to assume that the correlation between US bonds and US equities is the same for all bonds and equities
- And you are willing to assume that the correlation between US large-cap and small-cap stocks is the same for all stocks
- Then it is easy to calculate the covariance matrix for the entire bank
- Use perhaps DCC for US small caps and other portfolios
- And combine them by the constant correlation

Consider an Example

σ^2_{E} σ^2_{B} σ^2_{E} σ^2_{B}

where red is US, blue is Brazil, E denotes Equity and B Bond

Individual covariance matrices of your securities can be combined into a covariance matrix for the entire financial institution:



Advantages of the Hierarchical Approach

- Scalability — efficient even with large numbers of assets
- Dimensionality reduction — transforms multivariate modelling into a series of univariate problems
- Positive definiteness — ensured
- Interpretability — clear link between groups and risk sources
- Modularity — supports different models for different segments
- Ease of updating — local model changes do not require global recalibration

Limitations of the Hierarchical Approach

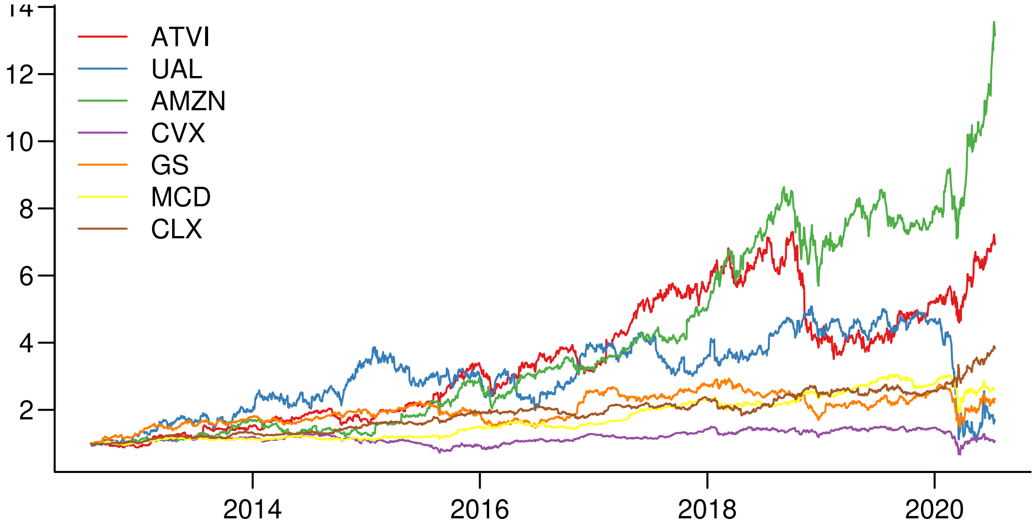
- Loss of nuance — group-level simplifications may mask important relationships
- Interpretation of components — transformed variables may lack intuitive meaning
- Model risk — strong structural assumptions may not hold in crises
- Static structure — if group definitions are fixed, model flexibility is constrained over time — grouping may become misaligned with real-world dynamics
- Sensitivity — the method can be affected by unusual observations or changes in the structure of the data over time
- Limited cross-group dynamics — interactions across segments may be underrepresented

Covid-19

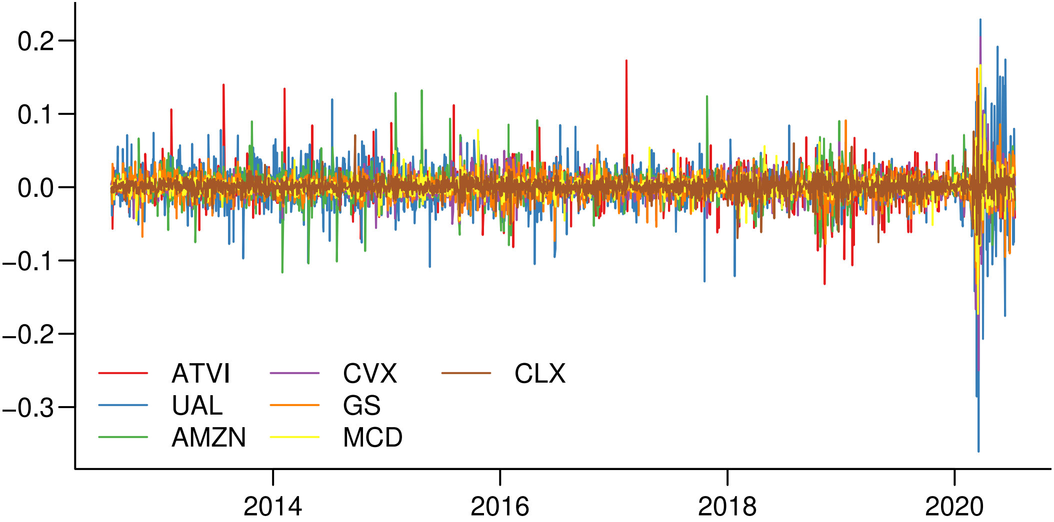
Multivariate Volatility Analysis

- Pick seven stocks
- Some that did quite well in the crisis
- Others that performed really badly
- And yet others that performed average
- Use DCC to get the conditional correlations
- Normalise the prices of all to start at one at the beginning of the sample

Normalised Prices (To Start at 1$)



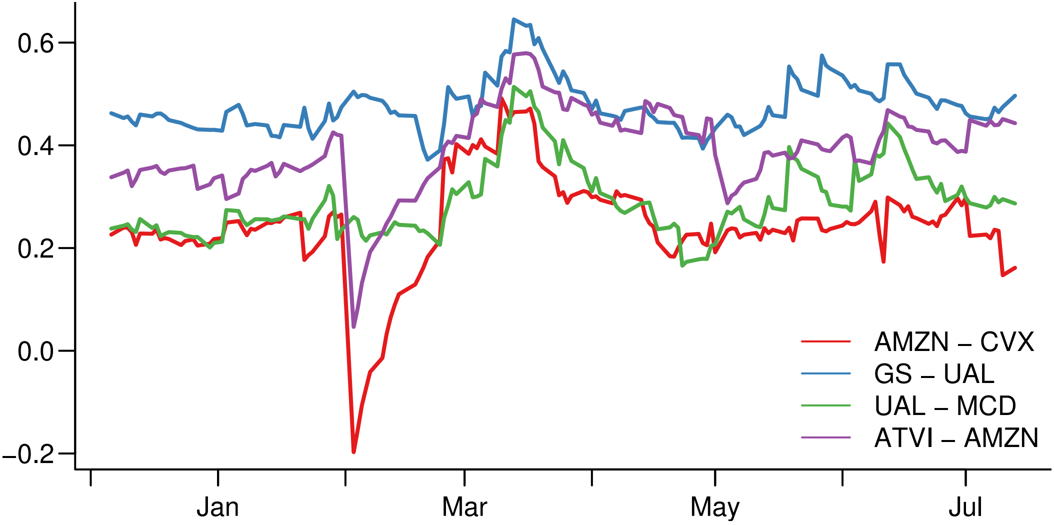
Returns


0.2
0.1
0.0
−0.1
−0.2
−0.3
2014
2016
2018
2020
ATVI
UAL
AMZN
CVX
GS
MCD
CLX

Amazon

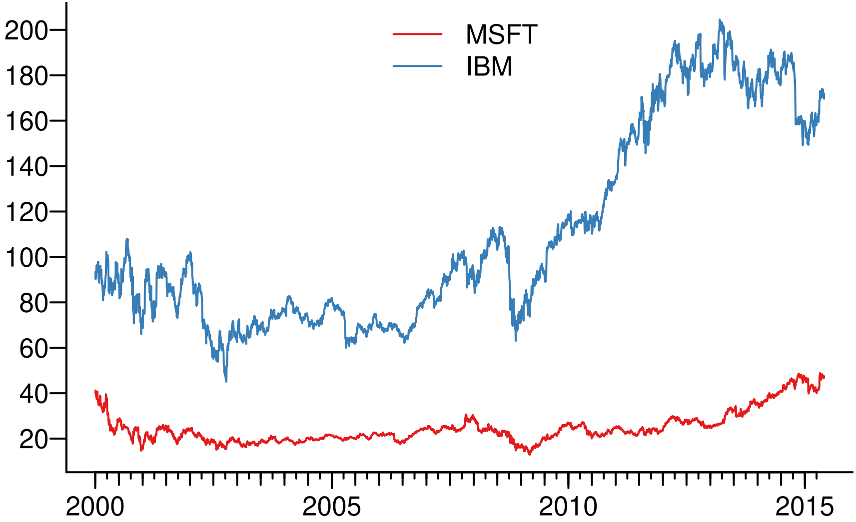
- Amazon prices jumped by 8.5% at the end of January
- Because of really good fourth-quarter results
- Note how that affects the conditional correlations

Pairwise Correlations

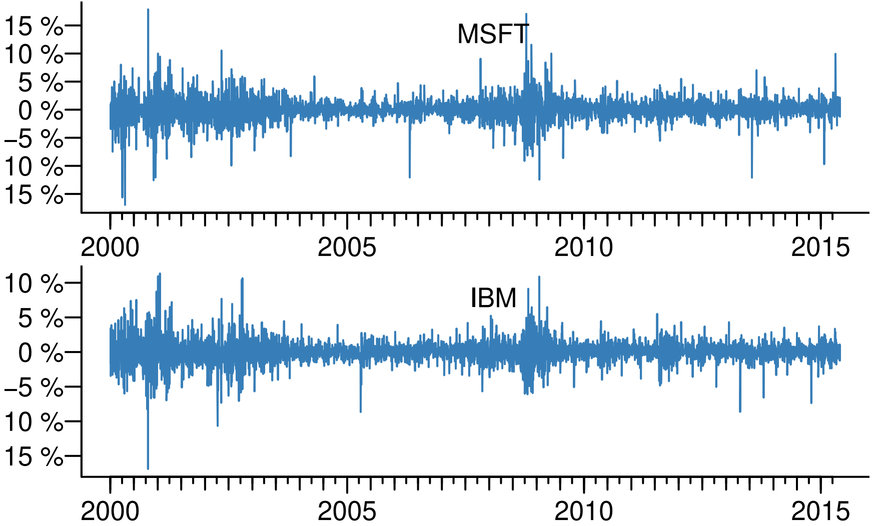


Estimation Comparison

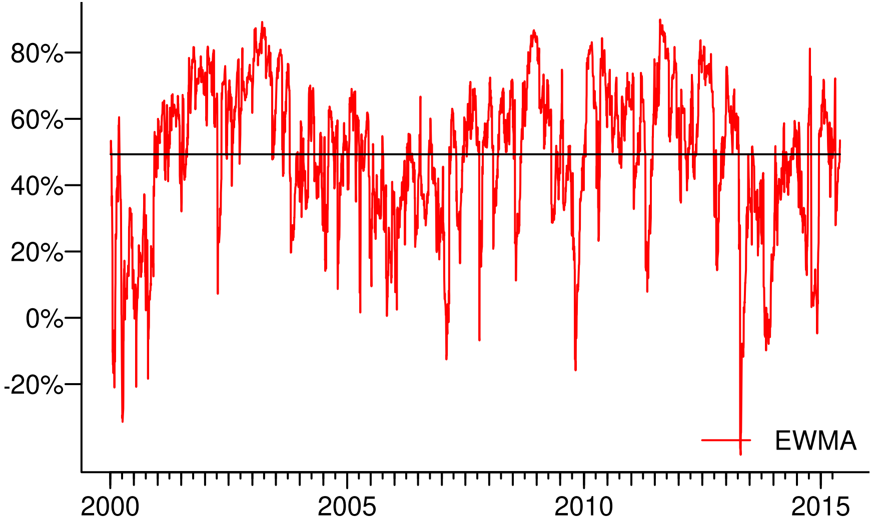
Prices



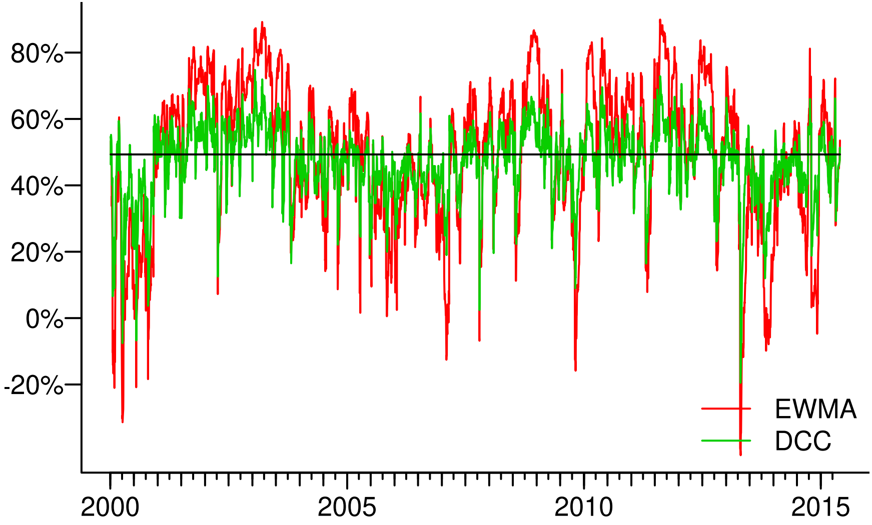
(B) Returns


MSFT
15 %
10 %
5 %
0 %
–5 %
10 %
15 %
2000
2005
2010
2015
IBM
10 %
5 %
0 %
–5 %
10 %
15 %
2000
2005
2010
2015

Correlation Estimates With Average Correlation 49%

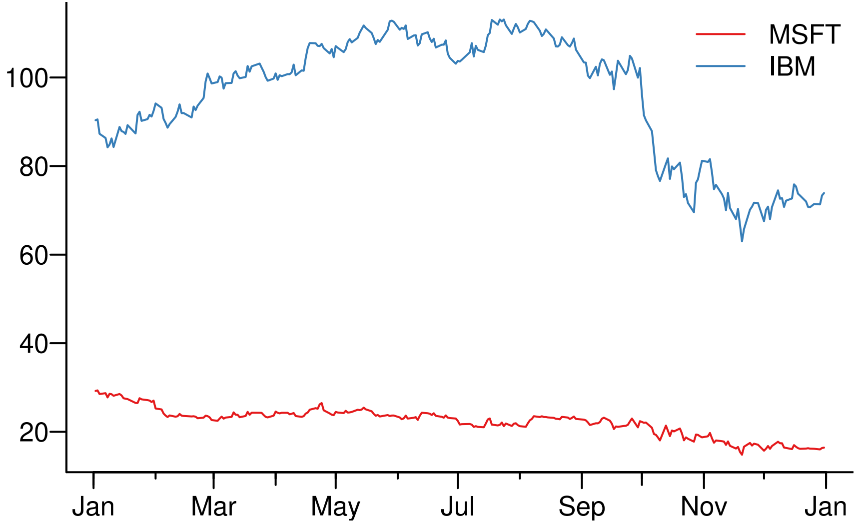


Correlation Estimates With Average Correlation 49%

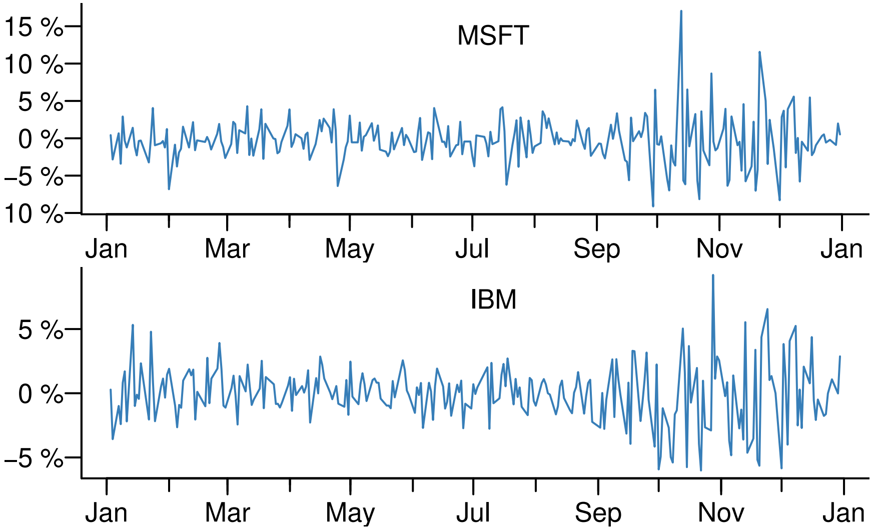


- Let's focus on 2008, the midst of the 2007 to 2009 crisis, when the correlations of all stocks increased dramatically

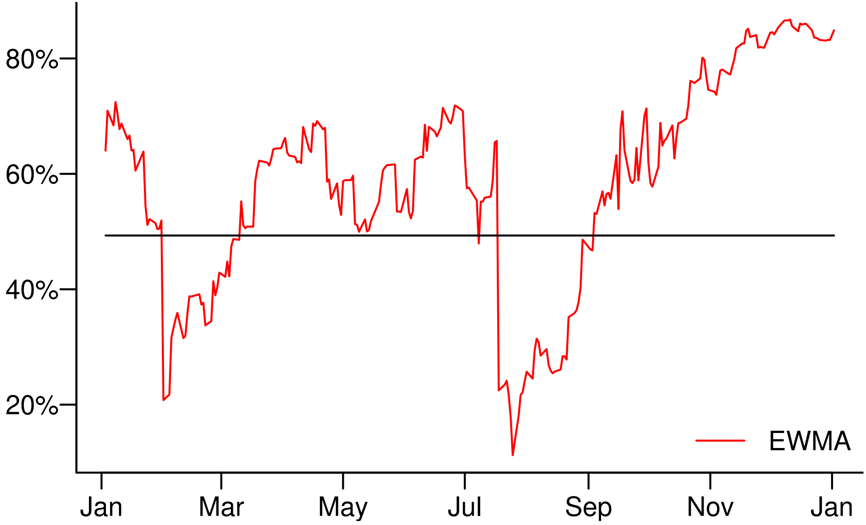
Prices



Returns


MSFT
15 %
10 %
5 %
0 %
–5 %
10 %
Jan
Mar
May
Jul
Sep
Nov
Jan
IBM
5 %
0 %
–5 %
Jan
Mar
May
Jul
Sep
Nov
Jan

Correlations, With Average Correlation 49%



Correlations, With Average Correlation 49%