Multivariate Volatility, Dependence and Copulas

Kevin Sheppard
http://www.kevinsheppard.com

Oxford MFE

This version: March 1, 2019

March 7, 2019

Overview

- Multivariate Volatility
  - Simple Models
  - Dynamic Models
- Realized Covariance
- Dependence
  - Linear (correlation)
  - Non-linear
- Copulas

Covariance Estimators

Principal Component Analysis

- Split returns in the orthogonal (uncorrelated) components

$$\min_{\boldsymbol{\beta},\mathbf{F}} (kT)^{-1} \sum_{i=1}^{k} \sum_{t=1}^{T} (y_{i,t} - \mathbf{f}_t \boldsymbol{\beta}_i)^2 \text{ subject to } \boldsymbol{\beta}'\boldsymbol{\beta} = \mathbf{I}_k$$

- Solution depends on eigenvalues and eigenvectors, easy to calculate
- Can order factors so that the partial R^2 are decreasing
- Factor 1 explains more than factor 2 which explains more than factor 3, etc.
- Can estimate the number of factors which are common across all assets
- *More details in notes ...*

Principal Component Covariance

Definition (n-period Principal Component Covariance)

The n-period principal component covariance is defined as

$$\boldsymbol{\Sigma}_t = \boldsymbol{\beta}_t' \boldsymbol{\Sigma}_t^f \boldsymbol{\beta}_t + \boldsymbol{\Omega}_t$$

where $\boldsymbol{\Sigma}_t^f = n^{-1} \sum_{i=1}^{n} \mathbf{f}_{t-i} \mathbf{f}_{t-i}'$ is the n-period moving covariance of first p principal component factors, $\hat{\boldsymbol{\beta}}_t$ is the p by k matrix of principal component loadings corresponding to the first p factors, and $\boldsymbol{\Omega}_t$ is a diagonal matrix with $\omega_{j,t+1}^2 = n^{-1} \sum_{i=1}^{n} \eta_{j,t-1}^2$ on the j$^{\text{th}}$ diagonal where $\eta_{i,t} = r_{i,t} - \mathbf{f}_t' \boldsymbol{\beta}_{i,t}$ are the residuals from a p-factor principal component analysis.

- Use principal components in place of observable factors
- Same advantage as observable factor covariance
  - Additional advantage that do not need factors
- Disadvantage that not as easy to implement in a structured setting
- Identical to moving average covariance if all factors used

Equicorrelation

- Assume that all correlations are identical

$$\sigma_{ij,t} = \rho\sigma_{i,t}\sigma_{j,t}$$

Definition (n-period Moving Average Equicorrelation Covariance)

The n-period moving average equicorrelation covariance is defined as

$$\boldsymbol{\Sigma}_t = \begin{bmatrix} \sigma_{1,t}^2 & \rho_t\sigma_{1,t}\sigma_{2,t} & \rho_t\sigma_{1,t}\sigma_{3,t} & \ldots & \rho_t\sigma_{1,t}\sigma_{k,t} \\ \rho_t\sigma_{1,t}\sigma_{2,t} & \sigma_{2,t}^2 & \rho_t\sigma_{2,t}\sigma_{3,t} & \ldots & \rho_t\sigma_{2,t}\sigma_{k,t} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_t\sigma_{1,t}\sigma_{k,t} & \rho_t\sigma_{2,t}\sigma_{k,t} & \rho_t\sigma_{3,t}\sigma_{k,t} & \ldots & \sigma_{k,t}^2 \end{bmatrix}$$

where $\sigma_{j,t}^2 = n^{-1}\sum_{i=1}^{n}\epsilon_{j,t}^2$ and ρ_t is estimated using one of the estimators below.

Equicorrelation

- Moment or maximum likelihood estimator for ρ
- Moment:

$$\mathrm{E}[\epsilon_{p,t}^2] = k^{-2}\sum_{j=1}^{k}\sigma_{j,t}^2 + 2k^{-2}\sum_{o=1}^{k}\sum_{q=o+1}^{k}\rho\sigma_{o,t}\sigma_{q,t}$$
$$= k^{-2}\sum_{j=1}^{k}\sigma_{j,t}^2 + 2\rho k^{-2}\sum_{o=1}^{k}\sum_{q=o+1}^{k}\sigma_{o,t}\sigma_{q,t}$$

- Estimator exploits this structure

$$\rho_t = \frac{\sigma_{p,t}^2 - k^{-2}\sum_{j=1}^{k}\sigma_{j,t}^2}{2k^{-2}\sum_{o=1}^{k}\sum_{q=o+1}^{k}\sigma_{o,t}\sigma_{q,t}}.$$

- $\sigma_{j,t}^2$ are the volatilities of the individual assets
- Only appropriate for homogeneous portfolios

Factor Models on the S&P 500

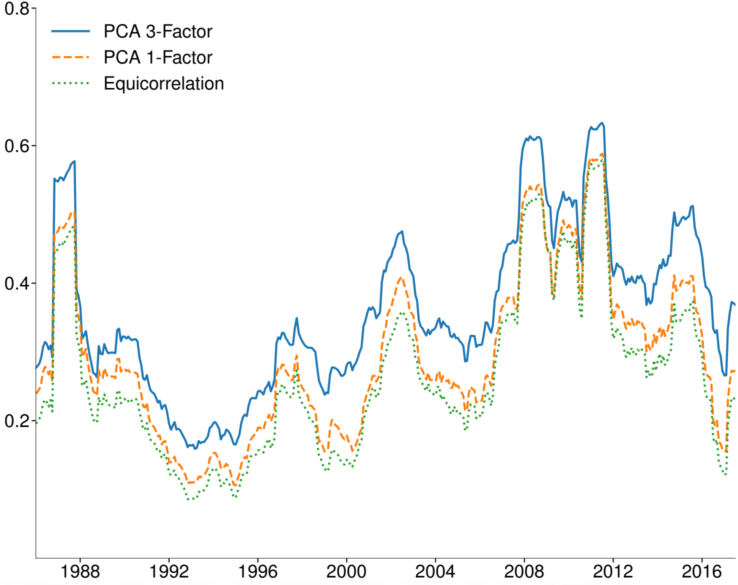
- Daily data on S&P 500 constituents from January 1, 1999 – December 31, 2008
- Return only included if present in relevant sample
  - Full sample
  - Rolling 252-day sample, centered at sample mid-point
- Full Sample PCA

k = 194	1	2	3	4	5	6
Partial R^2	0.263	0.039	0.031	0.023	0.019	0.016
Cumulative R^2	0.263	0.302	0.333	0.356	0.375	0.391

- Full Sample Correlation

Equicorrelation	1-Factor R^2 (S&P 500)	3-Factor R^2 (Fama-French)
0.255	0.236	0.267

Rolling Window Correlations



Dynamic Covariance Models

Exponentially Weighted Moving Average Covariance

- Simplest dynamic covariance model

Definition (Exponentially Weighted Moving Average Covariance)

The EWMA covariance is defined recursively as

$$\boldsymbol{\Sigma}_t = (1-\lambda)\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}'_{t-1} + \lambda\boldsymbol{\Sigma}_{t-1}$$

for $\lambda \in (0,1)$. EWMA covariance is equivalently defined through the infinite moving average

$$\boldsymbol{\Sigma}_t = (1-\lambda)\sum_{i=1}^{\infty}\lambda^{i-1}\boldsymbol{\epsilon}_{t-i}\boldsymbol{\epsilon}'_{t-i}.$$

RiskMetrics 1994 and 2006

- RiskMetrics popularized the use of EWMAs
- $\lambda = .94$ for daily data and $\lambda = .97$ for monthly
- RiskMetrics usually refers to the EWMA
- In 2006 it was updated to capture "long memory"

Definition (RiskMetrics 2006 Covariance)

The RiskMetrics 2006 Covariance is computed as

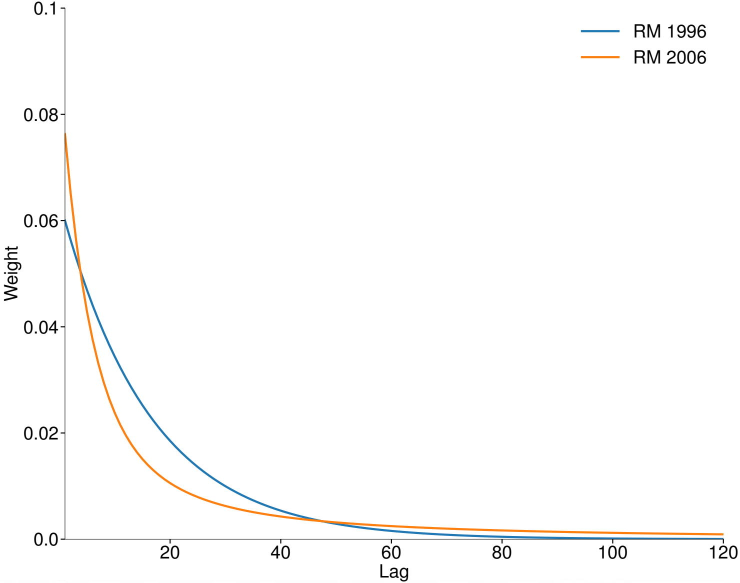
$$\boldsymbol{\Sigma}_t = \sum_{i=1}^{m} w_i \boldsymbol{\Sigma}_{i,t}$$

$$\boldsymbol{\Sigma}_{i,t} = (1-\lambda_i)\,\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}' + \lambda_i \boldsymbol{\Sigma}_{i,t-1}$$

$$w_i = \frac{1}{C}\left(1 - \frac{\ln(\tau_i)}{\ln(\tau_0)}\right),\ \lambda_i = \exp\left(-\frac{1}{\tau_i}\right),\ \tau_i = \tau_1 \rho^{i-1},\ \ i = 1, 2, \ldots, m$$

where C is a normalization constant which ensures that $\sum_{i=1}^{m} w_i = 1$.

The difference between RiskMetrics 1994 and 2006



Multivariate ARCH Models

- Like univariate models many specification
- Most specifications are interested in the trade-off between easy-of-estimation and parsimony
- Primary challenges are in making sure the conditional covariance is positive definite
  - More limiting that simple positivity int the scalar case.

Vector GARCH *vec*

- Most natural extension of GARCH
- Each cross-product drives each series
- Virtually impossible to use in practice
- Diagonal *vec* is more reasonable
- Usually use Matrix GARCH parameterization when using diagonal *vec*

BEKK (Baba, Engle, Kraft and Kroner)

- Parameterizes covariance in terms of positive (semi) definite components
- Assures covariance is positive definite.

Definition (BEKK GARCH)

The covariance in a BEKK GARCH(1,1) model evolves according to

$$\mathbf{\Sigma}_t = \mathbf{C}\mathbf{C}' + \mathbf{A}\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}'_{t-1}\mathbf{A}' + \mathbf{B}\mathbf{\Sigma}_{t-1}\mathbf{B}'$$

where $\mathbf{C}$ is a k by k lower triangular matrix and $\mathbf{A}$ and $\mathbf{B}$ are k by k parameter matrices.

- BEKK is a restricted *vec*
  - Does not work well when k is large
  - Number of parameters grows rapidly
- In bivariate model,

$$\sigma_{11,t} = c_{11}^2 + a_{11}^2\epsilon_{1,t-1}^2 + 2a_{11}a_{12}\epsilon_{1,t-1}\epsilon_{2,t-1} + a_{12}^2\epsilon_{2,t-1}^2$$
$$+b_{11}^2\sigma_{11,t-1} + 2b_{11}b_{12}\sigma_{12,t-1} + b_{12}^2\sigma_{22,t-1}.$$

Scalar and Diagonal BEKK

- Regular BEKK can be further restricted to reduce the number of parameters

Definition (Diagonal BEKK GARCH)

The covariance in a diagonal BEKK-GARCH(1,1) model evolves according to

$$\mathbf{\Sigma}_t = \mathbf{C}\mathbf{C}' + \tilde{\mathbf{A}}\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}'\tilde{\mathbf{A}}' + \tilde{\mathbf{B}}\mathbf{\Sigma}_{t-1}\tilde{\mathbf{B}}'.$$

where $\mathbf{C}$ is a k by k lower triangular matrix and $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ are diagonal parameter matrices.

- Can also use a scalar version

Definition (Scalar BEKK GARCH)

The covariance in a scalar BEKK-GARCH(1,1) model evolves according to

$$\mathbf{\Sigma}_t = \mathbf{C}\mathbf{C}' + a^2\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1} + b^2\mathbf{\Sigma}_{t-1}$$

where $\mathbf{C}$ is a k by k lower triangular matrix and a and b are scalar parameters.

Covariance Targeting Scalar BEKK

- Scalar BEKK can be covariance targeted
- Replaces the intercept ($\mathbf{C}\mathbf{C}'$) with a consistent estimator

$$(1 - a^2 - b^2)\overline{\boldsymbol{\Sigma}}$$

- $\overline{\boldsymbol{\Sigma}}$ is the long-run variance of the data.
- $\overline{\boldsymbol{\Sigma}}$ is usually estimated using the outer product of returns

$$\widehat{\overline{\boldsymbol{\Sigma}}} = T^{-1} \sum_{t=1}^{T} \boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t'$$

.

$$\boldsymbol{\Sigma}_t = (1 - a^2 - b^2)\widehat{\overline{\boldsymbol{\Sigma}}} + a^2 \boldsymbol{\epsilon}_{t-1} \boldsymbol{\epsilon}_{t-1} + b^2 \boldsymbol{\Sigma}_{t-1}$$

- Only a and b to be estimated using maximum likelihood
- Work when k is very large

Constant Conditional Correlation GARCH

- GARCH models for each variable, constant correlation
- Decomposes the conditional covariance into std. deviation and correlation

$$\mathbf{\Sigma}_t = \mathbf{D}_t \mathbf{R} \mathbf{D}_t.$$

- Conditional variances are typically modeled using standard GARCH(1,1) models
- Conditional correlation is the correlation of the devolatilized residuals,

$$u_{i,t} = \epsilon_{i,t} / \sigma_{i,t}$$

- Simple to estimate correlation with usual estimator on devolatilized
- Estimable when k is large
- Constant correlation implausible for most assets when the horizon is large

Constant Conditional Correlation GARCH

Definition (Constant Conditional Correlation GARCH)

The covariance in a constant conditional correlation GARCH model evolves according to

$$\boldsymbol{\Sigma}_t = \begin{bmatrix} \sigma_{11,t} & \rho_{12}\sigma_{1,t}\sigma_{2,t} & \rho_{13}\sigma_{1,t}\sigma_{3,t} & \cdots & \rho_{1k}\sigma_{1,t}\sigma_{k,t} \\ \rho_{12}\sigma_{1,t}\sigma_{2,t} & \sigma_{22,t} & \rho_{23}\sigma_{2,t}\sigma_{3,t} & \cdots & \rho_{2k}\sigma_{2,t}\sigma_{k,t} \\ \rho_{13}\sigma_{1,t}\sigma_{3,t} & \rho_{23}\sigma_{2,t}\sigma_{3,t} & \sigma_{33,t} & \cdots & \rho_{3k}\sigma_{3,t}\sigma_{k,t} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{1k}\sigma_{1,t}\sigma_{k,t} & \rho_{2k}\sigma_{2,t}\sigma_{k,t} & \rho_{3k}\sigma_{3,t}\sigma_{k,t} & \cdots & \sigma_{kk,t} \end{bmatrix}$$

where $\sigma^2_{i,t}$, $i = 1, 2, \ldots, k$ evolves according to some univariate GARCH process on asset i, usually a GARCH(1,1).

Dynamic Conditional Correlation GARCH

- Extends CCC to have scalar dynamics
- Volatilities follow GARCH dynamics (or other model)

Definition (Dynamic Conditional Correlation GARCH)

The covariance in a dynamic conditional correlation GARCH model evolves according to

$$\mathbf{\Sigma}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t.$$

where

$$\mathbf{R}_t = \mathbf{Q}_t \oslash \mathbf{Q}_t^*,$$

$$\mathbf{Q}_t = (1 - a - b)\,\overline{\mathbf{R}} + a\mathbf{u}_{t-1}\mathbf{u}_{t-1}' + b\mathbf{Q}_{t-1},$$

where $\mathbf{u}_t$ is the k by 1 vector of devolatilized returns ($u_{i,t} = \epsilon_{i,t}/\sqrt{\hat{\sigma}_{ii,t}}$)

- Nests CCC
- Mush more realistic, but may need to be richer when examining different asset classes

ARCH Models for Mutual Fund Covariance

- Examine multivariate GARCH models using 3 mutual funds
  - Oakmark I (OAKMX) - A broad large cap fund
  - Fidelity Small Cap Stock (FSLCX) - A broad small cap fund which seeks to invest in firms with capitalizations similar to those in the Russell 2000 or S&P 600.
  - Wasatch-Hoisington US Treasury (WHOSX) - A fund which invests at least 90% of total assets in U.S. Treasury securities and can vary the average duration of assets held from 1 to 25 years, depending on market conditions.
- Span 3 asset classes

ARCH Models for Mutual Fund Covariance

- CCC

	Large Cap	Small Cap	Bond
Large Cap	1	0.718	-0.258
Small Cap	0.718	1	-0.259
Bond	-0.258	-0.259	1

- Unconditional Correlation

	Large Cap	Small Cap	Bond
Large Cap	1	0.803	-0.306
Small Cap	0.803	1	-0.305
Bond	-0.306	-0.305	1

ARCH Models for Mutual Fund Covariance

- Scalar Models

	α	γ	β
DCC	$\underset{(3.4)}{0.009}$	–	$\underset{(4.9)}{0.990}$
Scalar BEKK	$\underset{(143.0)}{0.062}$	–	$\underset{(89.6)}{0.918}$
Asym. Scalar BEKK	$\underset{(158.9)}{0.056}$	$\underset{(84.7)}{0.021}$	$\underset{(65.8)}{0.911}$

Realized Covariance

Realized Covariance

- Multivariate version of realized variance
- Assume prices followed a k-variate diffusion

$$d\mathbf{p}_t = \boldsymbol{\mu}_t dt + \boldsymbol{\Omega}_t d\mathbf{W}_t$$

- $\boldsymbol{\mu}_t$ is the instantaneous drift
- $\boldsymbol{\Sigma}_t = \boldsymbol{\Omega}_t \boldsymbol{\Omega}_t'$ is the instantaneous covariance
- $d\mathbf{W}_t$ is a k-variate Brownian motion
- Realized covariance estimates

$$\int_0^1 \boldsymbol{\Sigma}_s ds$$

- In the presence of jumps, realized covariance estimates quadratic covariance
- Computed using high-frequency returns

$$RC_t = \sum_{i=1}^{m} \mathbf{r}_{i,t} \mathbf{r}_{i,t}' = \left(\mathbf{p}_{i,t} - \mathbf{p}_{i-1,t}\right)\left(\mathbf{p}_{i,t} - \mathbf{p}_{i-1,t}\right)',$$

Limitations of Realized Covariance

- Two major issues:
  - Prices are contaminated by noise (e.g. bid-ask bounce).
  - Prices are not perfectly synchronized.
- Standard method is to sample relatively infrequently, e.g. 5 – 20 minutes
- Can use subsampling to improve
- For example, 1-minutes prices but need to use 10-minute returns
- Uses *all* possible 10-minute returns, not just non-overlapping ones

$$
\begin{aligned}
RC_t^{SS} &= \frac{m}{n(m-n+1)} \sum_{i=1}^{m-n+1} \sum_{j=1}^{n} \mathbf{r}_{i+j-1,t}\mathbf{r}'_{i+j-1,t} \\
&= \frac{1}{n} \sum_{j=1}^{n} \frac{m}{(m-n+1)} \sum_{i=1}^{m-n+1} \mathbf{r}_{i+j-1,t}\mathbf{r}'_{i+j-1,t} \\
&= \frac{1}{n} \sum_{j=1}^{n} \widetilde{RC}_{j,t},
\end{aligned}
$$

  - m returns using highest frequency, n highest frequency returns

Realized Correlation and Beta

Definition (Realized Correlation)

The realized correlation between two series is defined

$$RCorr = \frac{RC_{ij}}{\sqrt{RC_{ii}RC_{jj}}}$$

where RC_{ij} is the realized covariance between assets i and j and RC_{ii} and RC_{jj} are the realized variances of assets i and j, respectively.

Realized Beta

Definition (Realized Beta)

Suppose RC_t is a $k+1$ by $k+1$ realized covariance matrix for an asset and a set of observable factors where the asset is in position 1, so that the realized covariance can be partitioned

$$RC = \left[\begin{array}{cc} RV_i & RC'_{fi} \\ RC_{fi} & RC_{f,f} \end{array} \right]$$

where $RV_{i,i}$ is the realized variance of the asset being studied, RC_{if} is the $kby1$ vector of realized covariance between the asset and the factors, and RC_{ff} is the k by k covariance of the factors. The Realized Beta is defined

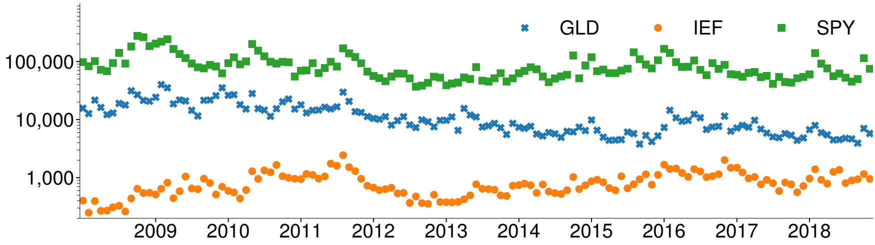
$$R\beta = RC_{ff}^{-1} RC_{fi}.$$

Application: ETF Realized Covariance

- Exchange-traded funds facilitate investing in assets that are often difficult to access for retail investors.
- Trade like stocks but backed by other assets
- Three ETFs:
  - SPDR S&P 500 ETF (SPY)
  - SPDR Gold Trust (GLD)
  - iShares 7-10 Year Treasury Bond ETF (IEF)

Transaction Counts

- Substantial heterogeneity in trading frequency
- SPY: 5+/second
- IEF: Between 30-120 seconds/trade



Determining the correct sampling frequency

Definition (Pseudo-Correlation Signature Plot)

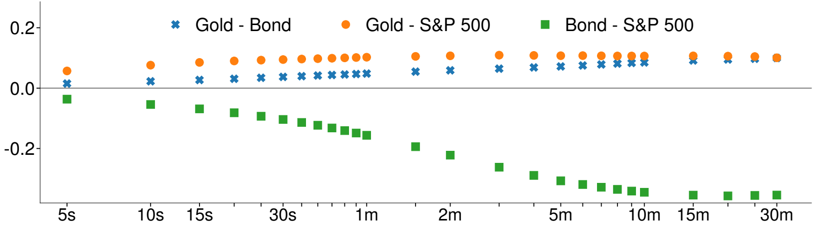
The pseudo-correlation signature plot displays the time-series average of Realized Covariance

$$\overline{RCorr}_{ij,t}^{(m)} = \frac{T^{-1}\sum_{t=1}^{T} RC_{ij,t}^{(m)}}{\overline{RVol}_i \times \overline{RVol}_j}$$

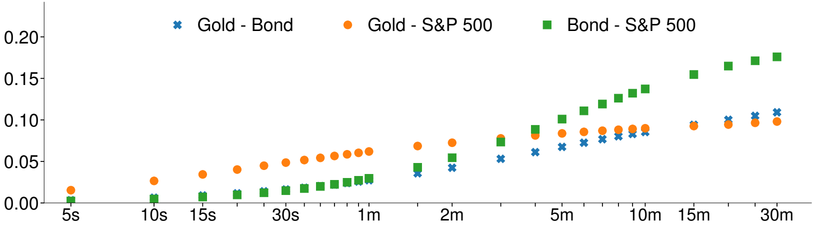
where m is the number of samples and $\overline{RVol}_{\bullet} = \sqrt{T^{-1}\sum_{t=1}^{T} RC_{\bullet\bullet,t}^{(q)}}$ is a the square root of the average Realized Variance using q-samples. q is chosen to produce an accurate RV that is free from microstructure effects. An equivalent representation displays the amount of time, whether in calendar time or tick time (number of trades between observations) along the X-axis.

Signature Plots

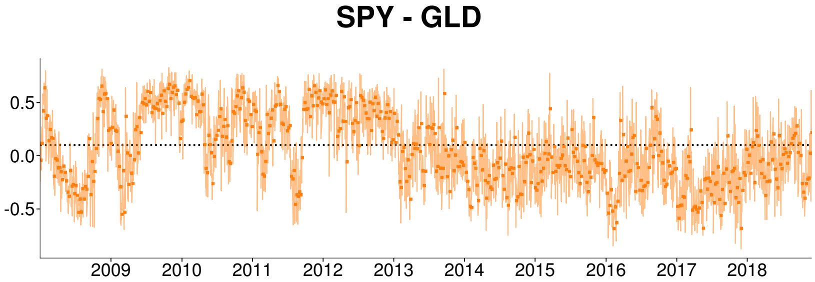
Pseudo Correlation Signature Plot

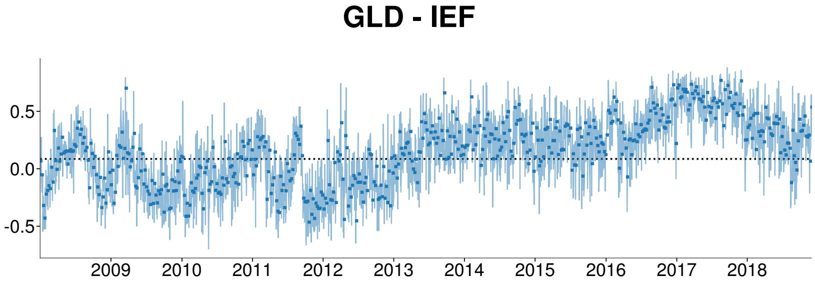


R^2 Signature Plot



Realized Correlations





Nonlinear Dependence

Beyond Linear Correlation

- Correlation and Beta both measure linear dependence
- Correlation is only closed under affine transformations ($a + bX$).
- Correlation is not robust to non-affine increasing transformations
- Correlation doesn't differentiate based on sign

Rank Correlation

- Correlation of the rank (CDF)

Definition (Rank (Spearman) Correlation)

The rank (Spearman) correlation between two random variables X and Y is

$$\rho_s(X,Y) = \text{Corr}\left(F_X(X), F_Y(Y)\right) = \frac{\text{Cov}\left[F_X(X), F_Y(Y)\right]}{\sqrt{\text{V}\left[F_X(X)\right]\text{V}\left[F_Y(Y)\right]}} = 12\text{Cov}\left[F_X(X), F_Y(Y)\right]$$

where the final identity uses the fact that the variance of a uniform (0,1) is $\frac{1}{12}$.

- Sample analogue

$$\rho = \frac{\sum_{i=1}^{n}\left(\frac{r_{x,i}}{n+1} - \frac{1}{2}\right)\left(\frac{r_{y,i}}{n+1} - \frac{1}{2}\right)}{\sqrt{\sum_{i=1}^{n}\left(\frac{r_{x,i}}{n+1} - \frac{1}{2}\right)^2}\sqrt{\sum_{j=1}^{n}\left(\frac{r_{y,i}}{n+1} - \frac{1}{2}\right)^2}}$$

- Invariant to monotone increasing non-linear transformations

Kendall's τ

- Uses *concordance* to measure dependence

Definition (Concordant Pair)

The pairs of random variables (x_i, y_i) and (x_j, y_j) are concordant if $\text{sgn}(x_i - x_j) = \text{sgn}(y_i - y_j)$ where $\text{sgn}(\cdot)$ is the sign function which returns -1 for negative values, 0 for zero, and +1 for positive values (or equivalently $(x_i - x_j)(y_i - y_j) > 0$).

- If a pair is not concordant then it is *discordant*.

Definition (Kendall's τ)

Kendall τ is defined

$$\tau = \Pr(\text{sgn}(x_i - x_j) = \text{sgn}(y_i - y_j)) - \Pr(\text{sgn}(x_i - x_j) \neq \text{sgn}(y_i - y_j))$$

- Sample analogue uses sample probabilities
- Invariant to monotone increasing non-linear transformations

Exceedance Correlation

- Differentiates between positive and negative returns
- Related to Expected Shortfall, which is an exceedance mean

Definition (Exceedance Correlation)

The exceedance correlation at level κ is defined as

$$\rho^{+}(\kappa) = \mathrm{Corr}\left(x, y | x > \kappa, y > \kappa\right)$$
$$\rho^{-}(\kappa) = \mathrm{Corr}\left(x, y | x < -\kappa, y < -\kappa\right)$$

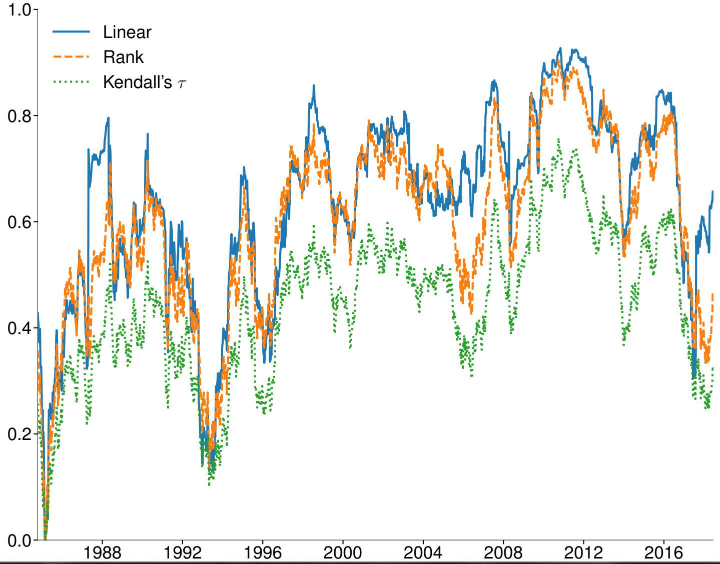
- Can use asset specific κs (e.g. κ_x and κ_y)
- Usual to use quantiles

Measuring Non-linear Dependence

- Daily data between S&P 500 and FTSE 100
- Data from 1984 until the present
- Use weekly returns to avoid synchronization issues
- Use rolling estimators, except for exceedance measures

Rolling Window Dependence

FTSE 100 and S&P 500



Exceedance Correlation

FTSE 100 and S&P 500