Lecture 9
Multivariate Volatility Modeling
Case study: Developing a Trading Strategy

Lars A. Lochstoer

UCLA Anderson School of Management

Winter 2020

Outline

1. Multivariate Volatility Models
   - MGARCH
     - CCC-GARCH
     - DCC-GARCH
2. Information Ratio
3. Case study
   - From two trading signals to full-fledged fund trading strategy

Multivariate Volatility Modeling

The Conditional Covariance Matrix

- Consider a vector of asset returns:

$$\mathbf{r}_t = \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t,$$

  where all variables are $N \times 1$ vectors

- The conditional means (modelled from, e.g., a VAR) are:

$$\boldsymbol{\mu}_t = E_{t-1}\left[\mathbf{r}_t\right]$$

- Denote the conditional $N \times N$ covariance matrix as:

$$\mathbf{H}_t = E_{t-1}\left[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'\right]$$

- A mean-variance optimizer would choose $N \times 1$ portfolio weights:

$$\mathbf{w}_{t-1} = k\mathbf{H}_t^{-1}\boldsymbol{\mu}_t$$

The Conditional Covariance Matrix (cont'd)

- Define the mean-zero, unit variance $N \times 1$ vector of i.i.d. shocks

$$\boldsymbol{\eta}_t \sim WN\left(\mathbf{0}_N, I_N\right).$$

This distribution is typically chosen to be Normal or Student's t

- We can then write:

$$\boldsymbol{\varepsilon}_t = \mathbf{H}_t^{1/2}\boldsymbol{\eta}_t$$

as

$$\begin{aligned}\boldsymbol{\varepsilon}_t &\sim WN\left(\mathbf{0}_N, \left(\mathbf{H}_t^{1/2}\right)' I_N \mathbf{H}_t^{1/2}\right) \\ &\sim WN\left(\mathbf{0}_N, \mathbf{H}_t\right)\end{aligned}$$

MGARCH

- Multivariate GARCH-type models take as their starting point

$$\mathbf{r}_t = \boldsymbol{\mu}_t + \mathbf{H}_t^{1/2} \boldsymbol{\eta}_t,$$

  where $\mathbf{H}_t^{1/2}$ is the Cholesky decomposition of $\mathbf{H}_t$

- These models specify the dynamics of $\mathbf{H}_t$
  - Needs to be positive definite for each t
  - Note: size of this matrix is of order N^2, so need to keep N small for good performance

Useful decomposition

We can write:

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t,$$

where the diagonal "standard deviations"-matrix is

$$\mathbf{D}_t = \begin{bmatrix} \sqrt{\sigma_{1,t}^2} & 0 & \cdots & 0 \\ 0 & \sqrt{\sigma_{2,t}^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\sigma_{N,t}^2} \end{bmatrix}$$

and the symmetric correlation matrix is

$$\mathbf{R}_t = \begin{bmatrix} 1 & \rho_{12,t} & \cdots & \rho_{1N,t} \\ \rho_{21,t} & 1 & \cdots & \rho_{2N,t} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{N1,t} & \rho_{N2,t} & \cdots & 1 \end{bmatrix}$$

Are individual portfolio StDevs and Corrs time-varying?

Consider monthly data from 1970 to 2017 for the Fama-French HML (value) and MOM (momentum) factors

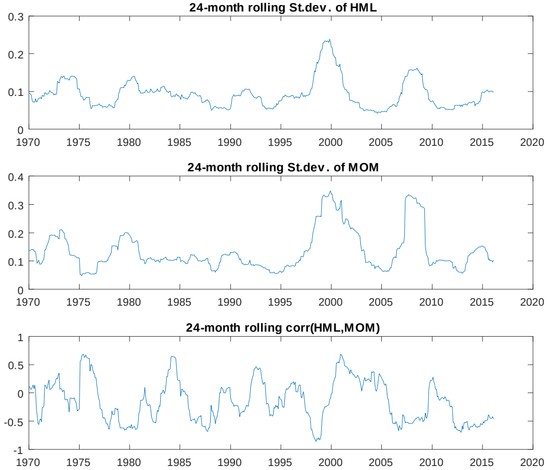
- Annualized summary statistics

	HML		MOM
$\mathbb{E}\, r^e_{t+1}$	0.039		0.064
σ	0.100		0.156
SR	0.383		0.411
ρ		-0.171	

Are individual portfolio StDevs and Corrs time-varying?

Plot 24-month rolling (annualized) st.devs. and correlation

- Yes! These moments look strongly time-varying
- MVE portfolio not likely to have constant portfolio weights



MGARCH - Specifications

- The diagonal elements of $\mathbf{D}_t$ can be obtained from N univariate GARCH models run on each element in $\mathbf{r}_t$
- The correlation matrix is tricker, in part as each correlation must be between -1 and 1 and the diagonal has to equal 1
- Two standard models:
  1. CCC-GARCH
     - ★ Constant conditional correlations
  2. DCC-GARCH
     - ★ Dynamic conditional correlations

CCC-GARCH

- Define:

$$\boldsymbol{\nu}_t = \mathbf{D}_t^{-1} \boldsymbol{\varepsilon}_t.$$

which implies that

$$\boldsymbol{\nu}_t \sim WN\left(0, \mathbf{R}_t\right).$$

- CCC-GARCH simply assumes that the correlation matrix is constant, thus:

$$\mathbf{R}_t = \mathbf{R} = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{\nu}_t \boldsymbol{\nu}_t'.$$

where each diagonal element of $\mathbf{D}_t$ is obtained from N univariate GARCH processes

DCC-GARCH(1,1)

- This model is due to Engle and Sheppard (2001), it is a benchmark model of time-varying conditional covariance matrix estimation
  - We still have problem if N is not small
- Main idea, put structure (AR(1)-like structure) on how conditional correlations move over time
  - Ensure well-behaved so positive definite conditional correlation matrix and each element between -1 and 1
- Decompose correlation matrix to achieve this:

$$
\begin{aligned}
\mathbf{R}_t &= \mathbf{Q}_t^{*-1}\mathbf{Q}_t\mathbf{Q}_t^{*-1}, \\
\mathbf{Q}_t &= (1-a-b)\,\overline{\mathbf{Q}} + a\boldsymbol{\nu}_{t-1}\boldsymbol{\nu}_{t-1}' + b\mathbf{Q}_{t-1} \\
\overline{\mathbf{Q}} &= \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{\nu}_t\boldsymbol{\nu}_t' \\
a &> 0, \quad b > 0, \quad a+b < 1
\end{aligned}
$$

DCC-GARCH (cont'd)

- Denote the ij'th element in $\mathbf{Q}_t$ as $q_{ij,t}$. Then:

$$\mathbf{Q}_t^* = \begin{bmatrix} \sqrt{q_{11,t}} & 0 & \cdots & 0 \\ 0 & \sqrt{q_{22,t}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{q_{NN,t}} \end{bmatrix}$$

- The likelihood function depends on the chosen distribution for the shocks (e.g., Normal), but is otherwise found in a way similar to the ARMA likelihood
  - Assume initial shocks equal unconditional average (in particular, need $\mathbf{Q}_0$ to be positive definite)
  - Use multivariate probability density function
  - This is all coded up in R/MatLab, so not covered here, though we will implement this in mini-case to follow

Pre-amble to trading strategy example

The Academic "Information Ratio" (The Appraisal Ratio)

Information Ratio

The information ratio (IR) is a standard performance metric:

$$IR_i = \frac{E\left[R_i - R_{benchmark}\right]}{\sigma\left[R_i - R_{benchmark}\right]}.$$

Let $R_{i,t}$ be excess returns to fund i and $R_{benchmark} = \beta_i MKT_t$

- Estimate α_i and β_i from the usual regression

$$R_{i,t} = \alpha_i + \beta_i' F_t + \varepsilon_{i,t}.$$

Then:

$$IR_i = \frac{E\left[R_{i,t} - \beta_i MKT_t\right]}{\sigma\left[R_{i,t} - \beta_i MKT_t\right]} = \frac{\alpha_i}{\sigma\left(\varepsilon_{i,t}\right)}$$

Information Ratio and maximal Sharpe Ratio

Assume the fund's information ratio is 0.3 and the market Sharpe ratio is 0.4

- What is the maximal Sharpe ratio one could achieve by combining the fund and the market?

$$\begin{aligned} \max SR &= \sqrt{SR_{MKT}^2 + IR_{Fund}^2} \\ &= \sqrt{0.3^2 + 0.4^2} = 0.5. \end{aligned}$$

- Math will follow in future lecture
- Notice that if $IR \neq 0$ it is possible to increase Sharpe ratio
  - Typically, though, we can't short sell a fund...

This is why α is interesting

- It means you can improve your Sharpe ratio relative to your benchmark.
- This is valuable and can therefore justify higher fees

Mini-Case:

From trading signals to trading strategy

Initial idea

- You have an idea about how to choose stocks that outperform existing benchmark portfolios
- In particular, you use a combination of textual analysis, stock prices, and social media-based data to come up with two trading signals for each stock in your trading universe
  1. A valuation signal (of fundemental value): $val_{i,t}$
  2. A sentiment signal (of shorter-term trend): $trend_{i,t}$
- Our task is to:
  1. See if the signals have any information
  2. If so, implement an efficient portfolio strategy trading based on these signals
- Caveat: For this illustrative exercise, we only do in-sample testing

Unconditional returns to simple trading strategies

- A natural starting point is to sort into portfolios based on the signals
- One could sort into decile portfolios for each signal (20 portfolios in total), for instance, and look at average return for each decile portfolio
  - Valid approach
  - Can see if there is significant spread in average returns to portfolio 10 vs portfolio 1 for each signal
  - Can also spot nonlinear effects (perhaps U-shaped average return pattern across portfolios)
  - Later, we will also use Fama-MacBeth regressions to do this

Unconditional returns to simple trading strategies

- Assume that we have time-series of excess returns to two trading strategies: $\{HML_t\}_{t=1}^{T}$ and $\{MOM_t\}_{t=1}^{T}$
  - The value and momentum factors from Kenneth French's webpage
  - Sample average returns, standard deviation, and Sharpe ratio; all annualized
  - Also give sample correlation (ρ)

	HML		MOM
$\mathbb{E}\ r_{t+1}^{e}$	0.039		0.064
σ	0.100		0.156
SR	0.383		0.411
ρ		-0.171	

Constructing a simple joint trading strategy

- Recall the conditional mean-variance efficient portfolio:

$$\omega_t = k\Sigma_t^{-1}\mu_t,$$

  where ω_t is a 2×1 vector of weights on HML and MOM, Σ_t is the 2×2 conditional covariance matrix of HML and MOM, and μ_t is a 2×1 vector with the conditional expected returns to HML and MOM

  - k is a constant we can find by matching the unconditional volatility of the resulting portfolio to some desired level of overall risk

- In the following we compare performance of portfolio with constant weights (based on unconditional sample covariance matrix and average excess returns) to that of a strategy that uses time-varying weights per the above equation.

Constructing a simple joint trading strategy

- Use the unconditional moments to form the unconditional mean-variance efficient portfolio based on these two sub-portfolios

$$
\begin{aligned}
w &= k \begin{bmatrix} 0.1^2 & -0.171 \times 0.1 \times 0.156 \\ -0.171 \times 0.1 \times 0.156 & 0.156^2 \end{bmatrix}^{-1} \begin{bmatrix} 0.039 \\ 0.064 \end{bmatrix} \\
&= k \begin{bmatrix} 4.74 \\ 3.14 \end{bmatrix}
\end{aligned}
$$

Find k be setting portfolio variance to 0.15^2:

$$
\begin{aligned}
0.15^2 &= k^2 \begin{bmatrix} 4.74 \\ 3.14 \end{bmatrix}' \begin{bmatrix} 0.1^2 & -0.171 \times 0.0156 \\ -0.171 \times 0.0156 & 0.156^2 \end{bmatrix} \begin{bmatrix} 4.74 \\ 3.14 \end{bmatrix} \\
&= k^2 \times 0.385
\end{aligned}
$$

Constructing a simple joint trading strategy (cont'd)

- Solving for k, we have:

$$k = \sqrt{\frac{0.15^2}{0.385}} = 0.24$$

and so the per-period portfolio weights are:

$$w = 0.24 \times \begin{bmatrix} 4.74 \\ 3.14 \end{bmatrix} = \begin{bmatrix} 1.14 \\ 0.75 \end{bmatrix}$$

- So, for a \$10M portfolio, put \$11.4M in HML, \$7.5M in MOM
  - If HML and MOM are long-short, zero-investment portfolios, this means long \$11.4M in value financed by shorting the same amount in growth, and long \$7.5M in winners, financed by the same amount short in losers. Finally, put \$10M in risk-free rate.
  - Or, if HML and MOM are long-only versions, you need to borrow \$8.9M in the risk-free rate to finance these positions.
  - Maintain these portfolio weights by rebalancing each month

Constructing a simple joint trading strategy (cont'd)

- The expected return and Sharpe ratio of this baseline trading strategy are then:

$$
\begin{aligned}
E\left[R^{MVE}\right] &= \begin{bmatrix} 1.14 & 0.75 \end{bmatrix} \begin{bmatrix} 0.039 \\ 0.064 \end{bmatrix} = 9.25\%. \\
SR\left(R^{MVE}\right) &= \frac{9.25\%}{15\%} = 0.62
\end{aligned}
$$

- Let's see if we can improve on this baseline by attempting to estimate
  - The conditional expected returns
  - The conditional covariance matrix

Estimating conditional expected returns

- Our tool for this is forecasting regressions.
  - Need predictive variables
  - Thus, we need extra instruments/signals to implement conditional strategies
  - We will keep it simple and estimate a $VAR(1)$ on HML and MOM returns:

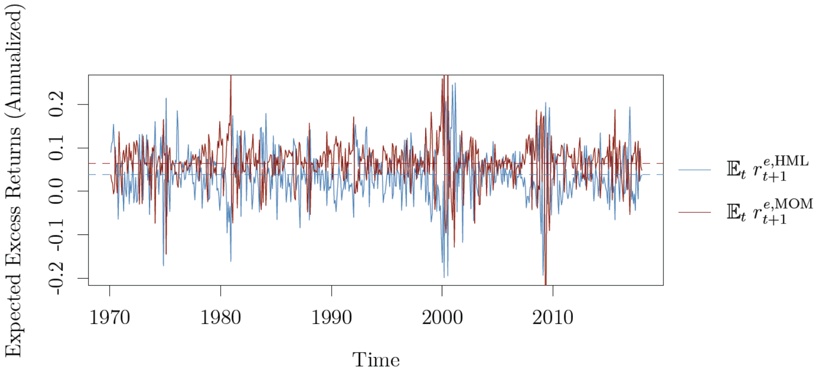
$$\begin{bmatrix} HML_t \\ MOM_t \end{bmatrix} = \begin{bmatrix} \phi_{01} \\ \phi_{02} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} HML_{t-1} \\ MOM_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{HML,t} \\ \varepsilon_{MOM,t} \end{bmatrix}$$

- Implementing this in R (see CCLE) , we get:

$$\begin{bmatrix} \hat{\phi}_{01} \\ \hat{\phi}_{02} \end{bmatrix} = \begin{bmatrix} 0.30 \\ 0.63 \end{bmatrix}, \begin{bmatrix} \hat{\phi}_{11} & \hat{\phi}_{12} \\ \hat{\phi}_{21} & \hat{\phi}_{22} \end{bmatrix} = \begin{bmatrix} 0.17 & -0.01 \\ -0.08 & 0.06 \end{bmatrix}, \begin{matrix} R^2_{HML} = 2.9\% \\ R^2_{MOM} = 0.7\% \end{matrix}$$

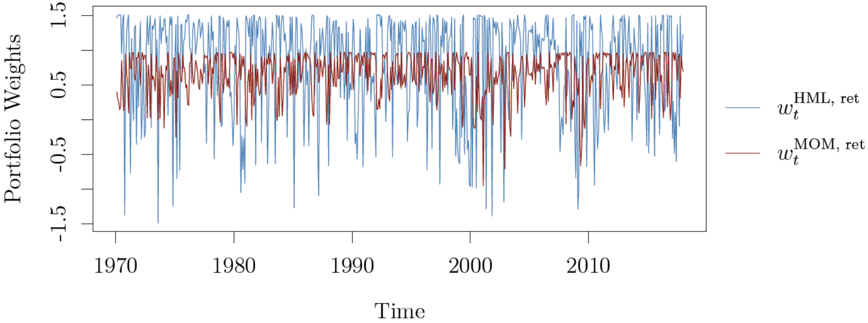
Expected Returns from the VAR(1)

- The time-series of annualized expected monthly returns is below
- Lots of variation, economically, even though R^2s were small



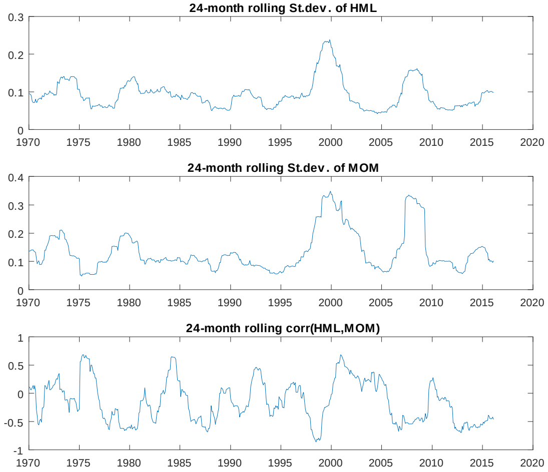
MVE Portfolio Weights

- Using the expected returns from the VAR, along with a constant covariance matrix of the residuals, we obtain portfolio weights
  - As before, set volatility of portfolio to be 15% annualized
  - HML was most predictable and therefore sees biggest changes in portfolio weights
  - Notice economically large changes in weights, despite modest R^2s in forecasting regressions
  - Full sample Sharpe ratio for this strategy is $0.78 > 0.61$ (from no expected return timing)



Adding Volatility Timing

- Recall that a cursory look at the data indicates substantial time-variation in the conditional covariance matrix



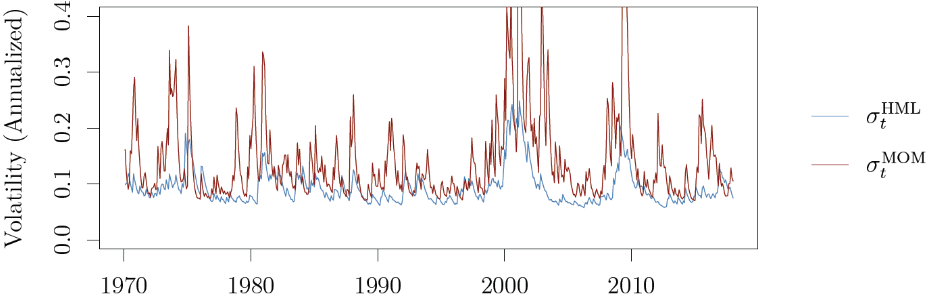
Estimating a DCC-GARCH(1,1)

- We will use the *ccgarch*-package in *R*

```
#Run GARCH on residuals

a <- c(0.003, 0.005)
A <- diag(c(0.2,0.3))
B <- diag(c(0.75, 0.6))
dcc.para <- c(0.01,0.98)

dcc.GARCH = dcc.estimation(inia = a, iniA = A, iniB = B, ini.dcc = dcc.para,
          dvar = data.dt[complete.cases(data.dt), .(eps.log.HML, eps.log.MOM)],
          model = "extended"
          )

data.dt[!is.na(shift(log.HML, 1)), `:=`(sigma.HML = sqrt(dcc.GARCH$h[, 1]),
                       sigma.MOM = sqrt(dcc.GARCH$h[, 2]),
                       rho = dcc.GARCH$DCC[, 2],
                       eta.HML = dcc.GARCH$std.resid[, 1],
                       eta.MOM = dcc.GARCH$std.resid[, 2]
                       )]
```

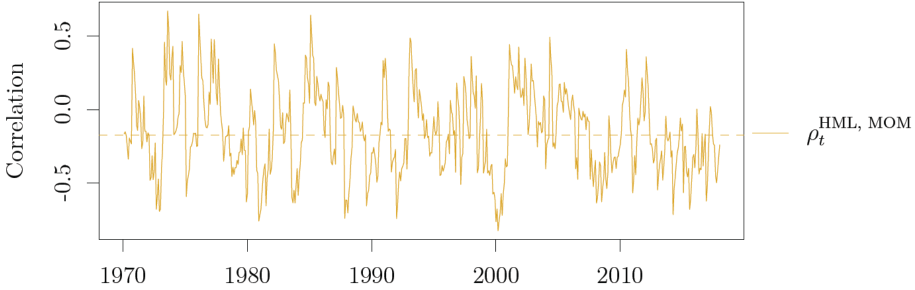
Estimated conditional volatilities

- Lots of time-variation, persistent processes
- Momentum has most time-variation in vol (least in expected return)



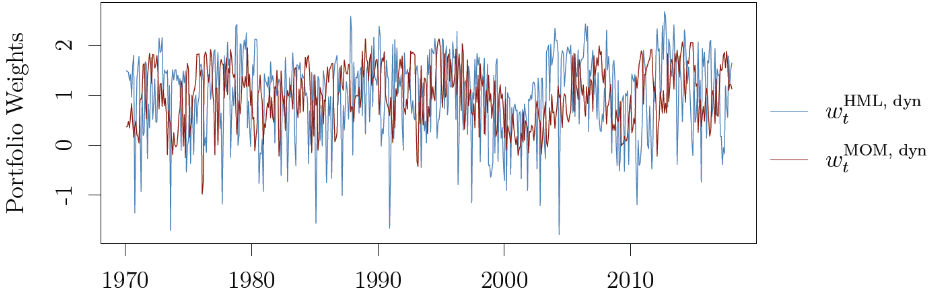
Estimated conditional correlation

- Conditional correlation between HML and MOM
- LOTS of time-variation here (-0.7 to 0.7), likely large effects on MVE weights



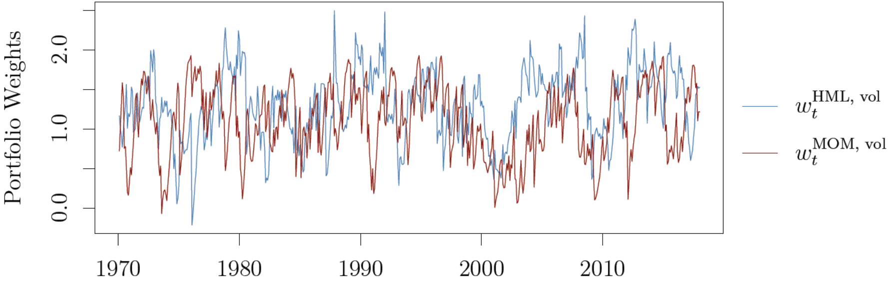
MVE weights from vol and return timing

- More variation than when only return timing
- The vol component is also more persistent (though this is likely a feature of the predictors for returns only being lagged return)



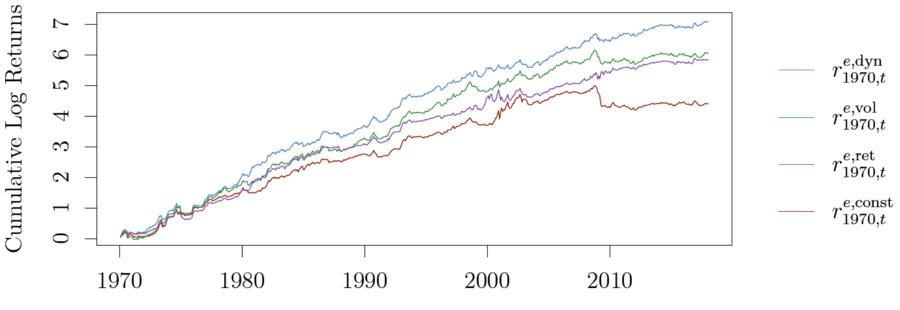
Additional strategy: Vol timing only

- Also run case where no attempt at forecasting return, only the conditional covariance matrix
- I.e., only vol timing. Less frequent trading, no shorting (pretty much).



Cumulative Log Returns for four strategies

- Strategies are: No timing, Er timing, Vol timing, both Er and Vol timing (dynamic)
- Ranking is: Both Er and Vol, only Vol, only Er, no timing
- Caveat: recall, these are in-sample tests



Summary statistics for four strategies

- Timing of return and variance-covariance matrix adds substantially to Sharpe ratios
  - Marginal SR increase: $IR = \sqrt{0.943^2 - 0.613^2} = 0.717$
  - Caveat: all in-sample and ignoring transaction costs
  - Trading is hazardous to your wealth

	Dynamic	Volatility-Timing	Return-Timing	Constant
$\mathbb{E}\, r^e_{t+1}$	0.148	0.127	0.122	0.092
σ	0.157	0.154	0.156	0.150
SR	0.943	0.824	0.781	0.613

	Correlation Matrix			
	Dynamic	Volatility-Timing	Return-Timing	Constant
Dyn.	1.000	0.816	0.873	0.667
Vol.		1.000	0.625	0.855
Ret.			1.000	0.592
Const.				1.000

On timing strategies

- In terms of MVE portfolio weights, the ones we calculated move too much
  - A lot of trading comes with a lot of trading costs (hazardous to your wealth)

- In practice, out-of-sample metrics lead to "shrinkage" of weights so they move less
  - Likely benefit is smaller than what we found in-sample
  - Market-timing is not easy
  - Neverthess, this should still be a part of your portfolio optimization