Econometric models & stochastic processes for prices, volatilities, spreads...
Number of factors explaining 55% of the variance versus VIX volatility index (2002-2008)
Number of EVs versus VIX (1/2006-2/2010)

- subprime
- Lehman, AIG, etc.
Dynamics are important

The previous slides show that the structure of the market is far from static.

This is obvious if we consider innovations in the market (new issues, new industries, the economic cycle, bubbles).

Equilibrium theories (e.g. APT, CAPM) are insufficient to explain prices, volatilities and correlations of financial assets.

Hence the need to model the evolution of financial variables using stochastic processes based on time-series analysis.

What can time-series analysis do for us?

-- Understand serial correlations in the data
-- Construct predictive models over suitable time-windows.
-- Discrete-time processes: important for data analysis.
-- Continuous-time processes: useful for theoretical purposes and to model high-dimensional data.
Stationarity/ Non Stationarity

**Definition:** a stochastic process is stationary if

$$\forall m, \; \forall (t_1, \ldots, t_m), \quad \forall E \in \mathbb{R}^n$$

$$\Pr\{\left(X_{t_1}, X_{t_2}, \ldots, X_{t_m}\right) \in A\} = \Pr\{\left(X_{t_1+h}, X_{t_2+h}, \ldots, X_{t_m+h}\right) \in A\}$$

A stationary process is a process that is **statistically invariant under translations**

Examples: the **Ornstein-Uhlembeck process** is stationary, Brownian motion is not.
The Ornstein-Uhlenbeck process

\[ dX_t = \kappa(m - X_t)dt + \sigma dW_t, \quad \kappa > 0 \]

\[ X_t = e^{-\kappa(t-s)}X_s + (1 - e^{-\kappa(t-s)})m + \sigma \int_s^t e^{-\kappa(t-u)}dW_u \]

\[ X_t = m + \sigma \int_{-\infty}^t e^{-\kappa(t-s)}\eta(s)ds, \quad \eta(s) = \text{Gaussian white noise} \]

Exponentially-weighted moving average of uncorrelated Gaussian random variables.
Serial correlations of the OU process

\[ \langle X_t X_{t+h} \rangle = \sigma^2 \left\langle \int_{-\infty}^{t} e^{-k(t-s)} \eta(s) ds \cdot \int_{-\infty}^{t+h} e^{-k(t+h-s')} \eta(s') ds' \right\rangle \]

\[ = \sigma^2 \int_{-\infty}^{t} \int_{-\infty}^{t+h} e^{-k(t-s)} e^{-k(t+h-s')} \delta(s-s') ds ds' \]

\[ = \sigma^2 \int_{-\infty}^{t} e^{-k(t-s)} e^{-k(t+h-s)} ds \]

\[ = \sigma^2 e^{-kh} \int_{-\infty}^{t} e^{-2k(t-s)} ds \]

\[ = \frac{\sigma^2 e^{-kh}}{2k} \]

\[ \langle |X_{t+h} - X_t|^2 \rangle = \frac{\sigma^2}{k} \left(1 - e^{-kh}\right) \]

Structure Function
Mean-reversion: a "quantitative" form of stationarity

\[ \langle |X_t - X_{t+h}|^2 \rangle = \frac{\sigma^2}{\kappa} (1 - e^{-\kappa h}) \]

\[ \kappa = \frac{1}{\tau}, \quad \tau = \text{characteristic time associated with correlation decay} \]
AR(1) model

\[ X_n = a + bX_{n-1} + \varepsilon_n \quad \varepsilon_n \sim N(0, \sigma^2) \]

\[ X_n = b^n X_0 + a \sum_{k=1}^{n} b^{n-k} + \sum_{k=1}^{n} b^{n-k} \varepsilon_k \]

\[ = b^n X_0 + a \frac{b^n - 1}{b - 1} + N\left(0, \sigma^2 \frac{b^{2n} - 1}{b^2 - 1}\right) \]

Stationarity: \( |b| < 1 \), \( \therefore \mu_{eq} = \frac{a}{1 - b} \), \( \sigma_{eq}^2 = \frac{\sigma^2}{1 - b^2} \)

Estimation of \( b \):

\[ b = \frac{\sum_{t=1}^{T} \left(X_{n-t} - \bar{X}\right) \left(X_{n-t-1} - \bar{X}\right)}{\sum_{t=1}^{T} \left(X_{n-t} - \bar{X}\right)^2} \quad (T = \text{time window}) \]
Estimation of AR(1) model

\[ \varepsilon_n = X_n - a - bX_{n-1} \quad \text{i.i.d. normals, } n = 0, \ldots, T \]

\[ \ln P = -\frac{1}{2\sigma^2} \sum_{n=1}^{T} (X_n - a - bX_{n-1})^2 - \frac{T}{2} \ln \sigma^2 - \frac{T}{2} \ln(2\pi) \]

\[ (a_{ml}, b_{ml}, \sigma_{ml}^2) = \arg \max_{a,b,\sigma^2} \ln P \quad \text{Maximum likelihood ~ minimum least squares} \]

\[ a_{ml} = \frac{\langle X_{n+1} \rangle \langle X_n^2 \rangle - \langle X_n X_{n+1} \rangle}{\langle X_n^2 \rangle - (\langle X_n \rangle)^2}, \quad b_{ml} = \frac{\langle X_n X_{n+1} \rangle - \langle X_n \rangle \langle X_{n+1} \rangle}{\langle X_n^2 \rangle - (\langle X_n \rangle)^2} \]

\[ \sigma_{ml}^2 = \langle (X_{n+1} - a_{ml} - b_{ml}X_n)^2 \rangle \]

where \[ \langle X_n \rangle = \frac{1}{T} \sum_{t=0}^{T-1} X_t, \quad \langle X_n \rangle = \frac{1}{T} \sum_{t=0}^{T-1} X_{t+1} \]
Estimation of Ornstein-Uhlenbeck models

\[ X_{t+\Delta t} = e^{-k\Delta t} X_t + m(1-e^{-k\Delta t}) + \sigma \int_{t}^{t+\Delta t} e^{-k(t-s)} dW_s \]

\[ X_{n+1} = a + bX_n + \varepsilon_{n+1} \{ \varepsilon_n \} \text{ i.i.d. } N\left(0, \sigma^2 \left(\frac{1-e^{-2k\Delta t}}{2k}\right)\right) \]

\[ b = \text{SLOPE}((X_{n-1},...,X_n);(X_{n-1},...,X_{n-1})) \]
\[ a = \text{INTERCEPT}((X_{n-1},...,X_n);(X_{n-1},...,X_{n-1})) \]
\[ k = \frac{1}{\Delta t} \ln \left( \frac{1}{b} \right), \quad m = \frac{a}{1-b}, \quad \sigma = \frac{\text{STDEV}(X_{n+1} - bX_n - a)}{\sqrt{1-b^2}} \sqrt{2 \frac{1}{\Delta t} \ln \left( \frac{1}{b} \right)} \]
Auto-regressive Models AR(m)

$X_1, X_2, \ldots, X_n, \ldots$ Time-series data to be modeled

$$X_n = a + \sum_{k=1}^{m} b_k X_{n-k} + \varepsilon_n \quad \varepsilon_n \sim N(0, \sigma^2), \text{ i.i.d.}$$

$$Y_n = A + BY_{n-1} + E_n$$

$$Y_n = B^n Y_0 + \sum_{k=1}^{n} B^{n-k} A + \sum_{k=1}^{n} B^{n-k} E_k$$

AR(m) is a "vector" AR(1) model
Stationarity of AR(m)

\[ \mu := E(X_n) \]

\[ E(X_n) = a + \sum_{k=1}^{m} b_k E(X_{n-k}) \quad \therefore \quad \mu = a + \mu \sum_{k=1}^{m} b_k \]

\[ \mu = \frac{a}{1 - \sum_{k=1}^{m} b_k} \quad \text{necessary condition for stationarity: } \sum_{k=1}^{m} b_k < 1 \]

\[ Z_n := X_n - \mu \quad Z_n \sim AR(m) \quad \text{with } a = 0 \]

\( B \) is a contraction iff all of its eigenvalues are less than 1

\[ \det(B - \lambda I) = (-1)^m \left( \lambda^m - \sum_{k=1}^{m} \lambda^{m-k} b_k \right) = (-1)^m P(\lambda) \]

All the roots of \( P(\lambda) \) must satisfy \(|\lambda| < 1\)
Auto-regressive Models
ARCH(p), GARCH(p,q)
Following R. Engle and T. Bollerslev
Conditional Mean and Conditional Variance

\[ y_t, \quad t = 1, 2, 3, ..., T \]

Given time series

\[ p(y_t \mid y_{t-1}, y_{t-2}, ...) = p(y_t \mid \Phi_{t-1}) \]

Model the conditional distributions

\[ y_t = \mu(\Phi_{t-1}) + \sigma(\Phi_{t-1})\varepsilon_t, \quad E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = 1 \]

Example: \( y_t \mid \Phi_{t-1} \sim N(\mu(\Phi_{t-1}), \sigma^2(\Phi_{t-1})) \)
Returns of S&P 500 Index
12/1/2000-2/26/2010
ARCH(p) (Engle, 1982)


\[ y_t = \alpha + \beta x_t + u_t \]

Uncorrelated residuals does not necessarily imply independent residuals

\[ u_t = h_t^{1/2} \varepsilon_t \quad E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = 1 \]

\[ h_t = a_0 + a_1 u_{t-1}^2 \]

Unlike in AR, the error is not assumed to have constant variance.

More generally,

\[ h_t = a_0 + \sum_{k=1}^{p} a_k u_{t-k}^2 \]

Conditional variance is a lagged sum of squared residuals, eg.

\[ h_t = \frac{1}{T} \sum_{k=1}^{T} u_{t-k}^2 \]
GARCH(p,q) (Bollerslev, 1986)

\[ u_t = h_t^{1/2} \varepsilon_t \quad \text{and} \quad E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = 1 \]

\[ h_t = \omega + \sum_{i=1}^{p} \alpha_i u_i^2 + \sum_{j=1}^{q} \beta_j h_{t-j} \]

Dependence on previous squared returns and previous conditional variances.

Most famous versions in practice: GARCH(1,1) or GARCH (1,p) which are basically AR(p) processes on the conditional variance driven by the squared-returns process.
GARCH(1,1)

\[ h_t = \omega + \alpha u_{t-1}^2 + \beta h_{t-1} \]

1-lag dependence

\[ h_t = \omega + \alpha u_{t-1}^2 + \beta(\omega + \alpha u_{t-2}^2 + \beta h_{t-2}) \]
\[ = \omega + \beta \omega + \alpha(\beta u_{t-2}^2 + u_{t-1}^2) + \beta^2 h_{t-2} \]
\[ \therefore \]

\[ h_t = \frac{\omega}{1-\beta} + \alpha \sum_{k=1}^{\infty} \beta^k u_{t-k}^2 \]

GARCH(1,1) is an exponentially weighted moving average of squared-errors. Beta determines the effective ``window size'' for estimation of conditional variance.
GARCH(1,2)

\[
\begin{pmatrix}
    h_t \\
    h_{t-1}
\end{pmatrix} = \begin{pmatrix}
    \omega \\
    0
\end{pmatrix} + \begin{pmatrix}
    \alpha & 0 \\
    0 & 0
\end{pmatrix} \begin{pmatrix}
    u_{t-1}^2 \\
    0
\end{pmatrix} + \begin{pmatrix}
    \beta_1 & \beta_2 \\
    1 & 0
\end{pmatrix} \begin{pmatrix}
    h_{t-1} \\
    h_{t-2}
\end{pmatrix}
\]

Stability condition: \[\lambda^2 - \beta_1 \lambda - \beta_2 = 0 \implies |\lambda| < 1\]

\[
h_t = \bar{h} + \sum_{k=1}^{\infty} \lambda_1^k u_{t-k}^2 + \sum_{k=1}^{\infty} \lambda_2^k u_{t-k}^2
\]

Intuitively, GARCH(1,2) is the sum of two EWMA with different time-scales (decay rates).

Notice however that the right-hand side depends on \( h \) as well, so the PDF of the conditional variance is not a chi-squared.

GARCH(1,p) is the sum of (at most) p EWMA.
Returns of S&P 500 Index
12/1/2000-2/26/2010
Fitting to GARCH(1,p)

We know that the tails of SPY are heavy and behave like Student $t$ with $df \sim 3.5$

This heavy-tailed behavior of stock prices can be modeled by assuming a static distribution (Student) or a time-dependent distribution with a GARCH-type stochastic conditional variance.

The latter approach (GARCH) has the advantage that it incorporates dynamics so it may capture "persistence" of volatility across time.

From a portfolio risk-management perspective, the situation is "cured" by assuming a Student-$t$ distribution with 3.5 degrees of freedom for returns (to capture tail behavior) and an EWMA variance which is adjusted daily to capture volatility clustering effects.

The question that remains is: what is the correct estimation window?
# GARCH(1,1) estimation of SPY returns

Method: ML - BFGS with analytical gradient  
date: 03-02-10  
time: 18:10  
Included observations: 2320  
Convergence achieved after 56 iterations

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>omega</td>
<td>2.85989E-06</td>
<td>3.9342E-07</td>
<td>7.269290633</td>
</tr>
<tr>
<td>alpha_1</td>
<td>0.698241421</td>
<td>0.020073908</td>
<td>34.78353205</td>
</tr>
<tr>
<td>beta_1</td>
<td>0.508888808</td>
<td>0.050794297</td>
<td>10.01862092</td>
</tr>
</tbody>
</table>

Log Likelihood: 7053.473574  
Jarque Bera: 12844.90612  
Ljung-Box: 65535
### GARCH(2,1) estimation

Method: ML - BFGS with analytical gradient  

**Date:** 03-03-10  
**Time:** 13:25  
**Included observations:** 2320  
**Convergence achieved after 45 iterations**  

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>omega</td>
<td>2.69557E-05</td>
<td>2.4E-06</td>
<td>11.25236</td>
<td>0</td>
</tr>
<tr>
<td>alpha_1</td>
<td>0.541398855</td>
<td>0.073788</td>
<td>7.337198</td>
<td>2.1805E-13</td>
</tr>
<tr>
<td>alpha_2</td>
<td>0.355438292</td>
<td>0.035892</td>
<td>9.90302</td>
<td>0</td>
</tr>
<tr>
<td>beta_1</td>
<td>0.268210539</td>
<td>0.045356</td>
<td>5.913404</td>
<td>3.3511E-09</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test</th>
<th>Value</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Likelihood</td>
<td>7060.668319</td>
<td></td>
</tr>
<tr>
<td>Jarque Bera</td>
<td>12844.90612</td>
<td>0</td>
</tr>
<tr>
<td>Ljung-Box</td>
<td>65535</td>
<td>65535</td>
</tr>
</tbody>
</table>
Garch(1,2)

Method: ML - BFGS with analytical gradient

date: 03-03-10
time: 13:34
Included observations: 2320
Convergence achieved after 54 iterations

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>omega</td>
<td>1.93253E-06</td>
<td>3.45079E-07</td>
<td>5.600257981</td>
<td>2.14033E-08</td>
</tr>
<tr>
<td>alpha_1</td>
<td>0.347594236</td>
<td>0.053959618</td>
<td>6.441747563</td>
<td>1.18106E-10</td>
</tr>
<tr>
<td>beta_1</td>
<td>0.417978993</td>
<td>0.040988575</td>
<td>10.19745117</td>
<td>0</td>
</tr>
<tr>
<td>beta_2</td>
<td>0.329591408</td>
<td>0.064169394</td>
<td>5.136271201</td>
<td>2.80243E-07</td>
</tr>
</tbody>
</table>

Log Likelihood 7119.174476
Jarque Bera 12844.90612 Prob 0
Ljung-Box 65535 Prob 65535
Which model should we use?

All three GARCH models fit the data very well, with high z-statistics.

Preference should be given to the model with smallest number of parameters, so GARCH(1,1) should be suitable.
Cointegration and Pairs Trading

In the previous lecture we saw some examples of pairs trading with ETFs

\[ X_t = \text{return on XLK} \]
\[ Y_t = \text{return on EBAY} \]

Perform \( m \) – day regression to construct residuals

\[ Y_t = \beta X_t + \varepsilon_t \]

\[ \beta = \text{SLOPE}((Y_{t-m},...,Y_{t-1}), (X_{t-m},...,X_{t-1})) \]

\[ \varepsilon_t = Y_t - \beta X_t \]

\[ P & L = 100 \prod_{k=1}^{t} (1 + \varepsilon_k) \]
\[ y_t = y_0 + \sum_{k=1}^{t} \ln(1 + \varepsilon_k) \]

Question of interest: is \( y_t \) stationary? Does \( y_t \) have a `unit root'?
Dickey-Fuller Test for Unit Roots (aka Augmented Dickey-Fuller test)

The Dickey-Fuller test is used to test for unit roots in statistical data.

Consider the following model for the differentiated time-series:

\[
\Delta y_t = \alpha + \beta t + \delta_0 y_{t-1} + \sum_{k=1}^{n} \delta_k \Delta y_{t-k} + \varepsilon_t, \quad \Delta y_t = y_t - y_{t-1}
\]

Null hypothesis: there is a unit root, i.e. \( \delta_0 = 0 \).

\[
DF = \frac{\hat{\delta}_0}{\text{stdev}(\hat{\delta}_0)}
\]

ADF Critical Values:
- Reject delta=0 if DF <
  - 1% level \(-3.970385\)
  - 5% level \(-3.415895\)
  - 10% level \(-3.130187\)

\( n \) is determined dynamically as part of the test (Akaike Information Criterion)
EBAY vs. XLK residuals

\[ Y_t = \text{daily return of EBAY} \]
\[ X_t = \text{daily return of XLK} \]

\[ \varepsilon_t = Y_t - \beta(t-1, t-60) \cdot X_t \]
\[ y_t = y_0 + \sum_{k=1}^{t} \ln(1+\varepsilon_k), \quad y_0 = 100 \]
Augmented DF test for EBAY/XLK

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>tseries(-1)</td>
<td>-0.025582</td>
<td>0.009132</td>
<td>-2.801401</td>
<td>0.005222</td>
</tr>
<tr>
<td>D(tseries(-1))</td>
<td>-0.104975</td>
<td>0.036984</td>
<td>-2.838362</td>
<td>0.004660</td>
</tr>
<tr>
<td>D(tseries(-2))</td>
<td>0.032844</td>
<td>0.037145</td>
<td>0.884206</td>
<td>0.376875</td>
</tr>
<tr>
<td>D(tseries(-3))</td>
<td>0.041696</td>
<td>0.036765</td>
<td>1.134124</td>
<td>0.257113</td>
</tr>
<tr>
<td>D(tseries(-4))</td>
<td>-0.139433</td>
<td>0.036498</td>
<td>-3.820306</td>
<td>0.000145</td>
</tr>
<tr>
<td>D(tseries(-5))</td>
<td>0.023322</td>
<td>0.036852</td>
<td>0.632844</td>
<td>0.527033</td>
</tr>
<tr>
<td>D(tseries(-6))</td>
<td>-0.103297</td>
<td>0.036384</td>
<td>-2.839106</td>
<td>0.004649</td>
</tr>
<tr>
<td>D(tseries(-7))</td>
<td>-0.123580</td>
<td>0.036566</td>
<td>-3.379630</td>
<td>0.000764</td>
</tr>
<tr>
<td>D(tseries(-8))</td>
<td>0.062589</td>
<td>0.036842</td>
<td>1.698850</td>
<td>0.089771</td>
</tr>
<tr>
<td>D(tseries(-9))</td>
<td>0.103669</td>
<td>0.036604</td>
<td>2.832135</td>
<td>0.004751</td>
</tr>
<tr>
<td>C</td>
<td>0.120657</td>
<td>0.043010</td>
<td>2.805296</td>
<td>0.005160</td>
</tr>
<tr>
<td>@trend</td>
<td>-0.000006</td>
<td>0.000003</td>
<td>-2.076142</td>
<td>0.038228</td>
</tr>
</tbody>
</table>

Best lag fit: 9

Cannot reject UR @ 90% level
EBAY vs. QQQQ residuals
Null Hypothesis: tseries has a unit root

Exogenous: Constant and linear Trend

Lag Length: 4 (Automatic Based on AIC, MAXLAG=10)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>tseries(-1)</td>
<td>-0.023280</td>
<td>0.008338</td>
<td>-2.791940</td>
<td>0.005374</td>
</tr>
<tr>
<td>D(tseries(-1))</td>
<td>-0.078624</td>
<td>0.036419</td>
<td>-2.158873</td>
<td>0.031179</td>
</tr>
<tr>
<td>D(tseries(-2))</td>
<td>0.019488</td>
<td>0.036533</td>
<td>0.533428</td>
<td>0.593897</td>
</tr>
<tr>
<td>D(tseries(-3))</td>
<td>0.030306</td>
<td>0.036525</td>
<td>0.829726</td>
<td>0.406960</td>
</tr>
<tr>
<td>D(tseries(-4))</td>
<td>-0.114959</td>
<td>0.036359</td>
<td>-3.161785</td>
<td>0.001632</td>
</tr>
<tr>
<td>C</td>
<td>0.109870</td>
<td>0.039251</td>
<td>2.799187</td>
<td>0.005256</td>
</tr>
<tr>
<td>@trend</td>
<td>-0.000002</td>
<td>0.000002</td>
<td>-0.918302</td>
<td>0.358759</td>
</tr>
</tbody>
</table>
ARMA(p,q) process

\[ y_t = a_0 + \sum_{k=1}^{p} a_k y_{t-k} + \sum_{l=1}^{q} b_l u_{t-k} + u_t \]

Combines autorregressive models with moving average models

Simple linear time-series model
timeseries: y
Method: Nonlinear Least Squares (Levenberg-Marquardt)
date: 03-03-10 time: 18:52
Included observations: 755
p = 1 - q = 1 - constant - manual selection

<table>
<thead>
<tr>
<th></th>
<th>Coefficient Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>4.627335411</td>
<td>0</td>
<td>148.9024</td>
</tr>
<tr>
<td>AR(1)</td>
<td>0.986154258</td>
<td>0</td>
<td>159.9401</td>
</tr>
<tr>
<td>MA(1)</td>
<td>-0.110605985</td>
<td>0</td>
<td>-2.998961</td>
</tr>
</tbody>
</table>

R-squared: 0.965239
Mean dependent var: 4.628068

Adjusted R-squared: 0.965147
S.D. dependent var: 0.071188

S.E. of regression: 0.013290
Akaike info criterion: -5.791955

Sum squared resid: 0.132821
Schwarz criterion: -5.773571

Log likelihood: 2189.462984
Durbin-Watson stat: 2.007356

Inverted AR-roots: 0.99
Inverted MA-roots: 0.11
Fitting y to an AR(1) process

timeseries: y

Method: Nonlinear Least Squares (Levenberg-Marquardt)

date: 03-03-10 time: 18:49

Included observations: 755

p = 1 - q = 0 - constant - manual selection

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>4.627528</td>
<td>0.03</td>
<td>168.4630632</td>
</tr>
<tr>
<td>AR(1)</td>
<td>0.98229241</td>
<td>0.01</td>
<td>143.6624447</td>
</tr>
</tbody>
</table>

R-squared: 0.964800
Mean dependent var: 4.628068

Adjusted R-squared: 0.964753
S.D. dependent var: 0.071188

S.E. of regression: 0.013365
Akaike info criterion: -5.782052

Sum squared resid: 0.134501
Schwarz criterion: -5.769796

Log likelihood: 2184.724802
Durbin-Watson stat: 2.225006
AR(1) coefficient for y estimated over a 60-day period

Red = upper bound for MR in 10 days, Green = upper bd for MR in 5 days
### Dickey-Fuller over Sep 2008/March 2009

Augmented Dickey-Fuller test statistic  
-2.593218  0.284178

Test critical values:

<table>
<thead>
<tr>
<th>Level</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1% level</td>
<td>-4.027516</td>
</tr>
<tr>
<td>5% level</td>
<td>-3.443485</td>
</tr>
<tr>
<td>10% level</td>
<td>-3.146482</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>tseries(-1)</td>
<td>-0.113728</td>
<td>0.043856</td>
<td>-2.593218</td>
<td>0.010671</td>
</tr>
<tr>
<td>D(tseries(-1))</td>
<td>-0.11532</td>
<td>0.090621</td>
<td>-1.230747</td>
<td>0.220785</td>
</tr>
<tr>
<td>D(tseries(-2))</td>
<td>0.162647</td>
<td>0.087448</td>
<td>1.859935</td>
<td>0.065303</td>
</tr>
<tr>
<td>D(tseries(-3))</td>
<td>0.040018</td>
<td>0.088750</td>
<td>0.450911</td>
<td>0.652854</td>
</tr>
<tr>
<td>D(tseries(-4))</td>
<td>-0.267631</td>
<td>0.085738</td>
<td>-3.121501</td>
<td>0.002247</td>
</tr>
<tr>
<td>D(tseries(-5))</td>
<td>0.076574</td>
<td>0.086639</td>
<td>0.883828</td>
<td>0.378528</td>
</tr>
<tr>
<td>D(tseries(-6))</td>
<td>-0.139433</td>
<td>0.085911</td>
<td>-1.623007</td>
<td>0.107169</td>
</tr>
<tr>
<td>D(tseries(-7))</td>
<td>-0.242743</td>
<td>0.082689</td>
<td>-2.935598</td>
<td>0.003980</td>
</tr>
<tr>
<td>D(tseries(-8))</td>
<td>0.090026</td>
<td>0.085786</td>
<td>1.049428</td>
<td>0.296056</td>
</tr>
<tr>
<td>D(tseries(-9))</td>
<td>0.189077</td>
<td>0.084225</td>
<td>2.244910</td>
<td>0.026575</td>
</tr>
<tr>
<td>D(tseries(-10))</td>
<td>-0.106441</td>
<td>0.084083</td>
<td>-1.265896</td>
<td>0.207962</td>
</tr>
<tr>
<td>C</td>
<td>0.511426</td>
<td>0.199365</td>
<td>2.565270</td>
<td>0.011521</td>
</tr>
<tr>
<td>@trend</td>
<td>0.000090</td>
<td>0.000044</td>
<td>2.058864</td>
<td>0.041636</td>
</tr>
</tbody>
</table>
AR-1 coefficient for the period Sep 2008/march 2009

timeseries: ebay/xlk
Method: Nonlinear Least Squares (Levenberg-Marquardt)
date: 03-03-10 time: 18:40
Included observations: 145
p = 1 - q = 0 - constant - manual selection

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>4.55489463</td>
<td>0.013841</td>
<td>329.0952386</td>
<td>0</td>
</tr>
<tr>
<td>AR(1)</td>
<td>0.88029659</td>
<td>0.035204</td>
<td>25.00582423</td>
<td>-2.2E-16</td>
</tr>
</tbody>
</table>

R-squared            | 0.813873     | Mean dependent var | 4.558974 |
Adjusted R-squared   | 0.812571     | S.D. dependent var | 0.045858 |
S.E. of regression   | 0.019853     | Akaike info criterion | -4.952615 |
Sum squared resid    | 0.056364     | Schwarz criterion | -4.911557 |
Log likelihood       | 361.064616   | Durbin-Watson stat | 2.312544 |
Conclusions

ARCH, GARCH: models for volatility of financial series.

Volatility analysis via ARCH and GARCH lead to exponential moving averages of squared returns.

The advantage of GARCH over a fixed window is that GARCH is endogenous. However, fixed estimation windows for volatilities and correlations or exogenous EWMAs also make sense from a risk-management perspective.

Cointegration of stock prices via pairs is not easy to establish econometrically.

Unit root test: tests for stationarity

ARMA, AR: models for mean-reversion
Mean-reversion & pairs trading
Systematic Approach for looking for mean-reversion in Equities

Look for stock returns devoid of explanatory factors, and analyze the corresponding residuals as *stochastic processes*.

\[ R_t = \sum_{k=1}^{m} \beta_k F_{kt} + \epsilon_t \]

Econometric factor model

\[ X_t = X_0 + \sum_{s=1}^{t} \epsilon_s \]

View residuals as increments of a process that will be estimated

\[ \frac{dS(t)}{S(t)} = \sum_{k=1}^{m} \beta_k \frac{dP_k(t)}{P_k(t)} + dX(t) \]

Continuous-time model for evolution of stock price
More on mean-reversion model

The factors are either

A. eigenportfolios corresponding to significant eigenvalues of the market

B. industry ETF, or portfolios of ETFs (we shall use these in light of last lecture and because it’s easier)

Questions of interest:

Can residuals be fitted to (increments of) OU processes or other MR processes?

If so, what is the typical correlation time-scale?

Experiment: consider 39 stocks associated with XLK (SPDR Tech ETF)
CSCO vs. XLK

![Graph comparing CSCO and XLK stock prices from 2009 to 2010](image-url)
Regressing returns of XYZ vs. XLK:
60-day window Betas (1/09-2/10)

<table>
<thead>
<tr>
<th></th>
<th>AAPL</th>
<th>ACS</th>
<th>ADBE</th>
<th>AKAM</th>
<th>APD</th>
<th>APH</th>
<th>BMC</th>
<th>CA</th>
<th>CPWR</th>
<th>CRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>1.03</td>
<td>0.66</td>
<td>1.35</td>
<td>1.3256</td>
<td>1.0962</td>
<td>1.28</td>
<td>0.72</td>
<td>1.04</td>
<td>1</td>
<td>1.46</td>
</tr>
<tr>
<td>Stdev</td>
<td>0.0959</td>
<td>0.16</td>
<td>0.14</td>
<td>0.2081</td>
<td>0.141</td>
<td>0.137</td>
<td>0.08</td>
<td>0.16</td>
<td>0.1</td>
<td>0.31</td>
</tr>
<tr>
<td>1pct q</td>
<td>0.8859</td>
<td>0.41</td>
<td>1.11</td>
<td>0.9823</td>
<td>0.8528</td>
<td>0.989</td>
<td>0.53</td>
<td>0.79</td>
<td>0.68</td>
<td>0.86</td>
</tr>
<tr>
<td>99 pct q</td>
<td>1.3376</td>
<td>1.02</td>
<td>1.66</td>
<td>1.7451</td>
<td>1.3665</td>
<td>1.571</td>
<td>0.89</td>
<td>1.34</td>
<td>1.14</td>
<td>1.94</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>CSCO</th>
<th>CTSH</th>
<th>CTXS</th>
<th>DELL</th>
<th>EBAY</th>
<th>EMC</th>
<th>ERTS</th>
<th>FISV</th>
<th>FLIR</th>
<th>GLW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>1.1764</td>
<td>1.08</td>
<td>1.13</td>
<td>1.2785</td>
<td>1.1911</td>
<td>1.1</td>
<td>1.08</td>
<td>0.9</td>
<td>1.01</td>
<td>1.36</td>
</tr>
<tr>
<td>Stdev</td>
<td>0.0512</td>
<td>0.12</td>
<td>0.13</td>
<td>0.1515</td>
<td>0.2107</td>
<td>0.065</td>
<td>0.11</td>
<td>0.11</td>
<td>0.13</td>
<td>0.14</td>
</tr>
<tr>
<td>1pct q</td>
<td>1.0743</td>
<td>0.84</td>
<td>0.84</td>
<td>1.0208</td>
<td>0.7202</td>
<td>0.944</td>
<td>0.88</td>
<td>0.7</td>
<td>0.67</td>
<td>1.13</td>
</tr>
<tr>
<td>99 pct q</td>
<td>1.2752</td>
<td>1.23</td>
<td>1.32</td>
<td>1.6897</td>
<td>1.5253</td>
<td>1.239</td>
<td>1.29</td>
<td>1.11</td>
<td>1.23</td>
<td>1.59</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>GOOG</th>
<th>HPQ</th>
<th>HRB</th>
<th>HRS</th>
<th>IBM</th>
<th>INTU</th>
<th>JDSU</th>
<th>JNPR</th>
<th>MFE</th>
<th>MSFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>0.7985</td>
<td>1.01</td>
<td>0.64</td>
<td>0.8516</td>
<td>0.7245</td>
<td>0.711</td>
<td>1.68</td>
<td>1.39</td>
<td>0.89</td>
<td>0.93</td>
</tr>
<tr>
<td>Stdev</td>
<td>0.1069</td>
<td>0.1</td>
<td>0.25</td>
<td>0.2352</td>
<td>0.0995</td>
<td>0.143</td>
<td>0.13</td>
<td>0.09</td>
<td>0.15</td>
<td>0.17</td>
</tr>
<tr>
<td>1pct q</td>
<td>0.644</td>
<td>0.86</td>
<td>0.2</td>
<td>0.4447</td>
<td>0.5158</td>
<td>0.472</td>
<td>1.41</td>
<td>1.16</td>
<td>0.59</td>
<td>0.64</td>
</tr>
<tr>
<td>99 pct q</td>
<td>1.0359</td>
<td>1.2</td>
<td>1.17</td>
<td>1.3077</td>
<td>0.9092</td>
<td>0.977</td>
<td>1.96</td>
<td>1.56</td>
<td>1.2</td>
<td>1.15</td>
</tr>
</tbody>
</table>
Regressing returns of XYZ vs. XLK: 60-day window Betas

<table>
<thead>
<tr>
<th></th>
<th>NOVL</th>
<th>NTAP</th>
<th>ORCL</th>
<th>QCOM</th>
<th>RHT</th>
<th>SRCL</th>
<th>SYMC</th>
<th>YHOO</th>
</tr>
</thead>
<tbody>
<tr>
<td>average</td>
<td>1.0213</td>
<td>0.66</td>
<td>1.34</td>
<td>1.3209</td>
<td>1.0915</td>
<td>1.273</td>
<td>0.72</td>
<td>1.05</td>
</tr>
<tr>
<td>stdev</td>
<td>0.1471</td>
<td>0.17</td>
<td>0.17</td>
<td>0.2288</td>
<td>0.1607</td>
<td>0.163</td>
<td>0.1</td>
<td>0.16</td>
</tr>
<tr>
<td>1pct q</td>
<td>0.3472</td>
<td>0.4</td>
<td>1.02</td>
<td>0.9302</td>
<td>0.7936</td>
<td>0.984</td>
<td>0.52</td>
<td>0.82</td>
</tr>
<tr>
<td>99 pct q</td>
<td>1.3399</td>
<td>1.02</td>
<td>1.66</td>
<td>1.7456</td>
<td>1.3667</td>
<td>1.572</td>
<td>0.89</td>
<td>1.34</td>
</tr>
</tbody>
</table>

Cross-sectional statistics for Beta variability

- min stdev 0.0512 CSCO
- max stdev 0.3142 CRM
- min range 0.2009 CSCO
- max range 1.0766 CRM
CRM vs. XLK
Evolution of 60-day Betas versus XLK: AAPL, CSCO, CRM, GOOG
Computing the residuals in practice

\[ X_1, \ldots, X_T \quad \text{ETF returns} \]
\[ Y_1, \ldots, Y_T \quad \text{stock returns} \]
\[ w = \text{estimation window (in days)} \]

\[ \beta_{t-w,t} = \text{SLOPE}((X_{t-w}, \ldots, X_{t-1}), (Y_{t-w}, \ldots, Y_{t-1})) \]

\[ \varepsilon_t = Y_t - \beta_{t-w,t} X_t, \quad t = w+1, w+2, \ldots, T \quad \text{Use window of } w \text{ days before current date} \]

Define \( Z_t := \sum_{k=w+1}^{t} \varepsilon_k \) \quad \text{``Co-integrated'' residual (CR)}
AAPL Residuals (against 60-day Betas)
CSCO residuals against 60-day Betas
Computing Mean-reversion from 10/30/09 to 1/28/10

Slope b is computed using lagged regression of CR

<table>
<thead>
<tr>
<th>Ticker</th>
<th>AAPL</th>
<th>ACS</th>
<th>ADBE</th>
<th>AKAM</th>
<th>APD</th>
<th>APH</th>
<th>BMC</th>
<th>CA</th>
<th>CPWR</th>
<th>CRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>b (slope)</td>
<td>0.95</td>
<td>0.98</td>
<td>0.88</td>
<td>0.87</td>
<td>0.94</td>
<td>0.86</td>
<td>0.76</td>
<td>0.89</td>
<td>0.97</td>
<td>0.85</td>
</tr>
<tr>
<td>kappa</td>
<td>11.9</td>
<td>4.43</td>
<td>30.9</td>
<td>35.3</td>
<td>15.7</td>
<td>36.9</td>
<td>67.7</td>
<td>28.3</td>
<td>8.89</td>
<td>41.7</td>
</tr>
<tr>
<td>tao (in days)</td>
<td>21.1</td>
<td>56.8</td>
<td>8.15</td>
<td>7.13</td>
<td>16</td>
<td>6.83</td>
<td>3.72</td>
<td>8.91</td>
<td>28.3</td>
<td>6.04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ticker</th>
<th>CSCO</th>
<th>CTSH</th>
<th>CTXS</th>
<th>DELL</th>
<th>EBAY</th>
<th>EMC</th>
<th>ERTS</th>
<th>FISV</th>
<th>FLIR</th>
<th>GLW</th>
</tr>
</thead>
<tbody>
<tr>
<td>b (slope)</td>
<td>0.93</td>
<td>0.76</td>
<td>1.02</td>
<td>0.93</td>
<td>0.8</td>
<td>0.82</td>
<td>0.89</td>
<td>0.9</td>
<td>0.92</td>
<td>0.97</td>
</tr>
<tr>
<td>kappa</td>
<td>17.3</td>
<td>68</td>
<td>-5.7</td>
<td>17.4</td>
<td>56.1</td>
<td>49.4</td>
<td>28.1</td>
<td>27.3</td>
<td>21.6</td>
<td>7.4</td>
</tr>
<tr>
<td>tao (in days)</td>
<td>14.6</td>
<td>3.71</td>
<td>-44.2</td>
<td>14.5</td>
<td>4.49</td>
<td>5.1</td>
<td>8.98</td>
<td>9.22</td>
<td>11.7</td>
<td>34.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ticker</th>
<th>GOOG</th>
<th>HPQ</th>
<th>HRB</th>
<th>HRS</th>
<th>IBM</th>
<th>INTU</th>
<th>JDSU</th>
<th>JNPR</th>
<th>MFE</th>
<th>MSFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>b (slope)</td>
<td>0.96</td>
<td>0.81</td>
<td>0.97</td>
<td>0.88</td>
<td>0.7</td>
<td>0.87</td>
<td>0.91</td>
<td>0.88</td>
<td>0.93</td>
<td>0.75</td>
</tr>
<tr>
<td>kappa</td>
<td>10.9</td>
<td>52.5</td>
<td>8.66</td>
<td>32.5</td>
<td>91.3</td>
<td>36.4</td>
<td>25.1</td>
<td>31.3</td>
<td>17.8</td>
<td>73.2</td>
</tr>
<tr>
<td>tao (in days)</td>
<td>23.1</td>
<td>4.8</td>
<td>29.1</td>
<td>7.75</td>
<td>2.76</td>
<td>6.93</td>
<td>10</td>
<td>8.04</td>
<td>14.2</td>
<td>3.44</td>
</tr>
</tbody>
</table>
Computing Mean-reversion from 10/30/09 to 1/28/10

<table>
<thead>
<tr>
<th>Ticker</th>
<th>NOVL</th>
<th>NTAP</th>
<th>ORCL</th>
<th>QCOM</th>
<th>RHT</th>
<th>SRCL</th>
<th>SYMC</th>
<th>YHOO</th>
</tr>
</thead>
<tbody>
<tr>
<td>b (slope)</td>
<td>0.98</td>
<td>0.9</td>
<td>0.96</td>
<td>0.78</td>
<td>0.88</td>
<td>0.91</td>
<td>0.88</td>
<td>0.91</td>
</tr>
<tr>
<td>kappa</td>
<td>5.48</td>
<td>25.8</td>
<td>9.28</td>
<td>61.7</td>
<td>33.4</td>
<td>22.9</td>
<td>31.2</td>
<td>25</td>
</tr>
<tr>
<td>tao(in days)</td>
<td>46</td>
<td>9.78</td>
<td>27.2</td>
<td>4.08</td>
<td>7.55</td>
<td>11</td>
<td>8.08</td>
<td>10.1</td>
</tr>
</tbody>
</table>
Structure function log (SLB/OIH)
Data: Apr 2006 to Feb 2009

OIH: Oil Services ETF, SLB: Schlumberger-Doll Research
Structure Function: long-short equal dollar weighted SLB-OIH

\[ P_{n+1} = P_n \times \left(1 + R_{\text{slb}} - R_{\text{oih}} \right), \quad X_n = \ln P_n \]
Structure Function for Beta-Neutral long-short portfolio SLB-Beta*OIH

\[ P_{n+1} = P_n \times \left(1 + R_{slb} - \beta_{60d} \cdot R_{oih}\right), \quad X_n = \ln P_n \]
Structure Function log (GENZ/IBB)

$S(x)$
Structure function ln (DNA/GENZ)

DNA: Genentech Inc.
GENZ; Genzyme Corp.

Mean-reversion: large negative curvature here.
Poor reversion for the beta adjusted pair. Beta is low ~ 0.30