Risk and Portfolio Management
Spring 2010

Auto-regressive Models

ARCH(p), GARCH(p,q)

Following R. Engle and T. Bollerslev

Conditional Mean and Conditional Variance

$y_t, \quad t = 1,2,3,\ldots,T$ Given time series

$p(y_t \mid y_{t-1}, y_{t-2}, \ldots) = p(y_t \mid \Phi_{t-1})$ Model the <u>conditional distributions</u>

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

Example: $\quad y_t \mid \Phi_{t-1} \sim N\left(\mu(\Phi_{t-1}), \sigma^2(\Phi_{t-1})\right)$

ARCH(p) (Engle, 1982)

$$y_t = \alpha + \beta x_t + u_t$$

Uncorrelated residuals does not necessarily imply independent residuals

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

$$h_t = a_0 + a_1 u_{t-1}^2$$

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)

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

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

Dependence on previous squared returns and previous conditional variances.

<u>Most famous versions in practice</u>: 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

$$h_t = \omega + \alpha u_{t-1}^2 + \beta\left(\omega + \alpha u_{t-2}^2 + \beta h_{t-2}\right)$$
$$= \omega + \beta\omega + \alpha\left(\beta u_{t-2}^2 + u_{t-1}^2\right) + \beta^2 h_{t-2}$$
$$\therefore$$

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

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)

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

Vector AR(1)

Stability condition: $\lambda^2 - \beta_1 \lambda - \beta_2 = 0 \Rightarrow |\lambda| < 1$

$$h_t = \bar{h} + A \sum_{k=1}^{\infty} \lambda_1^k u_{t-k}^2 + B \sum_{k=1}^{\infty} \lambda_2^k u_{t-k}^2$$

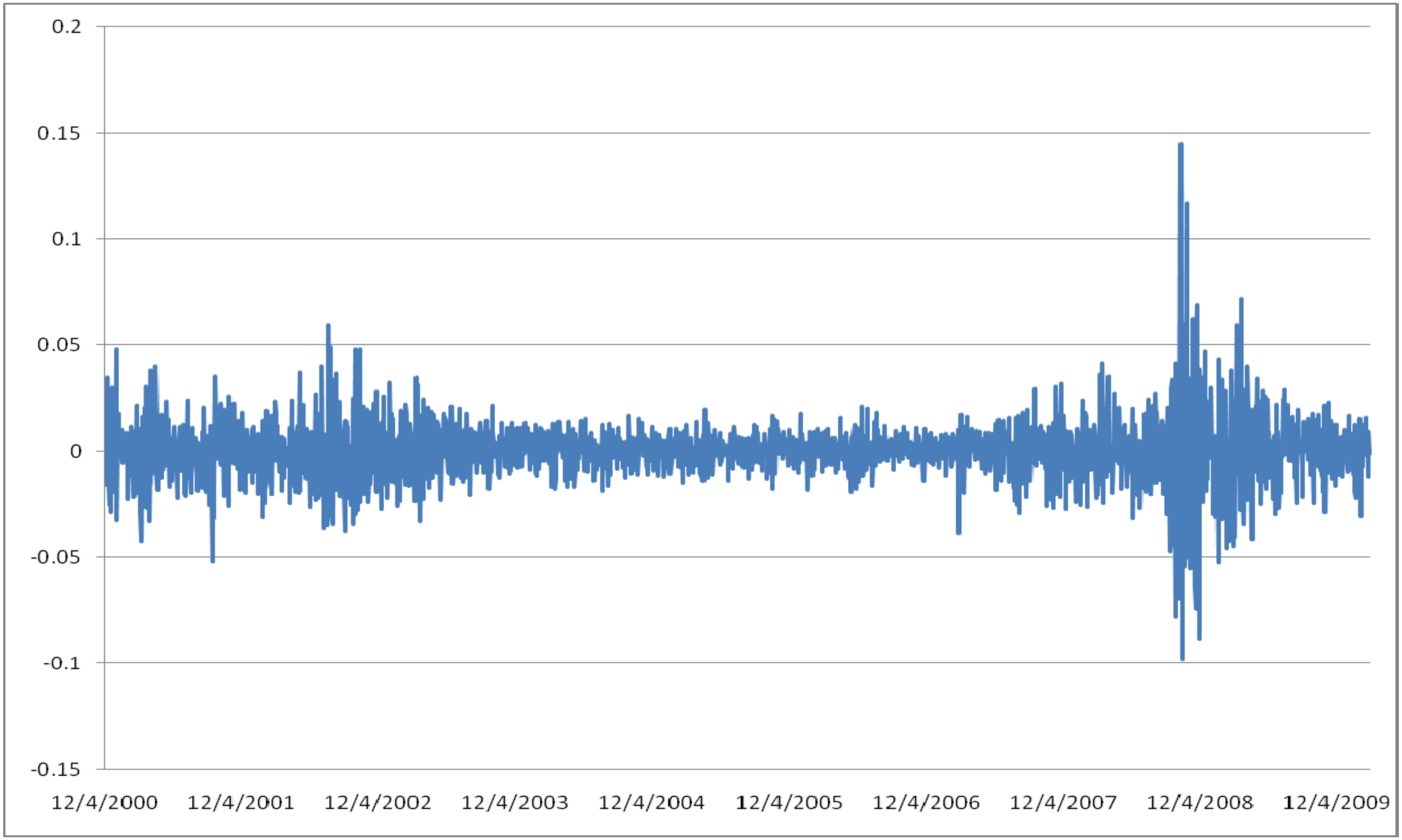
Steady-state solution

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 EWMAs.

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~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

	Coefficient	Std. Error	z-Statistic	Prob.
omega	2.85989E-06	3.9342E-07	7.269290633	3.61489E-13
alpha_1	0.698241421	0.020073908	34.78353205	0
beta_1	0.508888808	0.050794297	10.01862092	0

Log Likelihood	7053.473574			
Jarque Bera	12844.90612	Prob		0
Ljung-Box	65535	Prob		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

	Coefficient	Std. Error	z-Statistic	Prob.
omega	2.69557E-05	2.4E-06	11.25236	0
alpha_1	0.541398855	0.073788	7.337198	2.1805E-13
alpha_2	0.355438292	0.035892	9.90302	0
beta_1	0.268210539	0.045356	5.913404	3.3511E-09

Log Likelihood	7060.668319			
Jarque Bera	12844.90612	Prob		0
Ljung-Box	65535	Prob		65535

Garch(1,2)

Method: ML - BFGS with analytical gradient

date: 03-03-10

time: 13:34

Included observations: 2320

Convergence achieved after 54 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
omega	1.93253E-06	3.45079E-07	5.600257981	2.14033E-08
alpha_1	0.347594236	0.053959618	6.441747563	1.18106E-10
beta_1	0.417978993	0.040988575	10.19745117	0
beta_2	0.329591408	0.064169394	5.136271201	2.80243E-07

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 = return on XLK

Y_t = return on EBAY

Perform $m-$day regression to construct residuals

$$Y_t = \beta X_t + \varepsilon_t$$

$$\beta = \mathrm{SLOPE}\left(\left(Y_{t-m},\dots,Y_{t-1}\right),\left(X_{t-m},\dots,X_{t-1}\right)\right)$$

$$\varepsilon_t = Y_t - \beta X_t$$

$$\mathrm{P\,\&\,L} = 100 * \prod_{k=1}^{t}(1+\varepsilon_k) \qquad y_t = y_0 + \sum_{k=1}^{t}\ln(1+\varepsilon_k)$$

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:

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

Null hypothesis: there is a unit root, i.e. $\delta_0 = 0.$ $\quad DF = \frac{\hat{\delta}_0}{\mathrm{stdev}\left(\hat{\delta}_0\right)}$

n is determined dynamically
as part of the test
(Akaike Information Criterion)

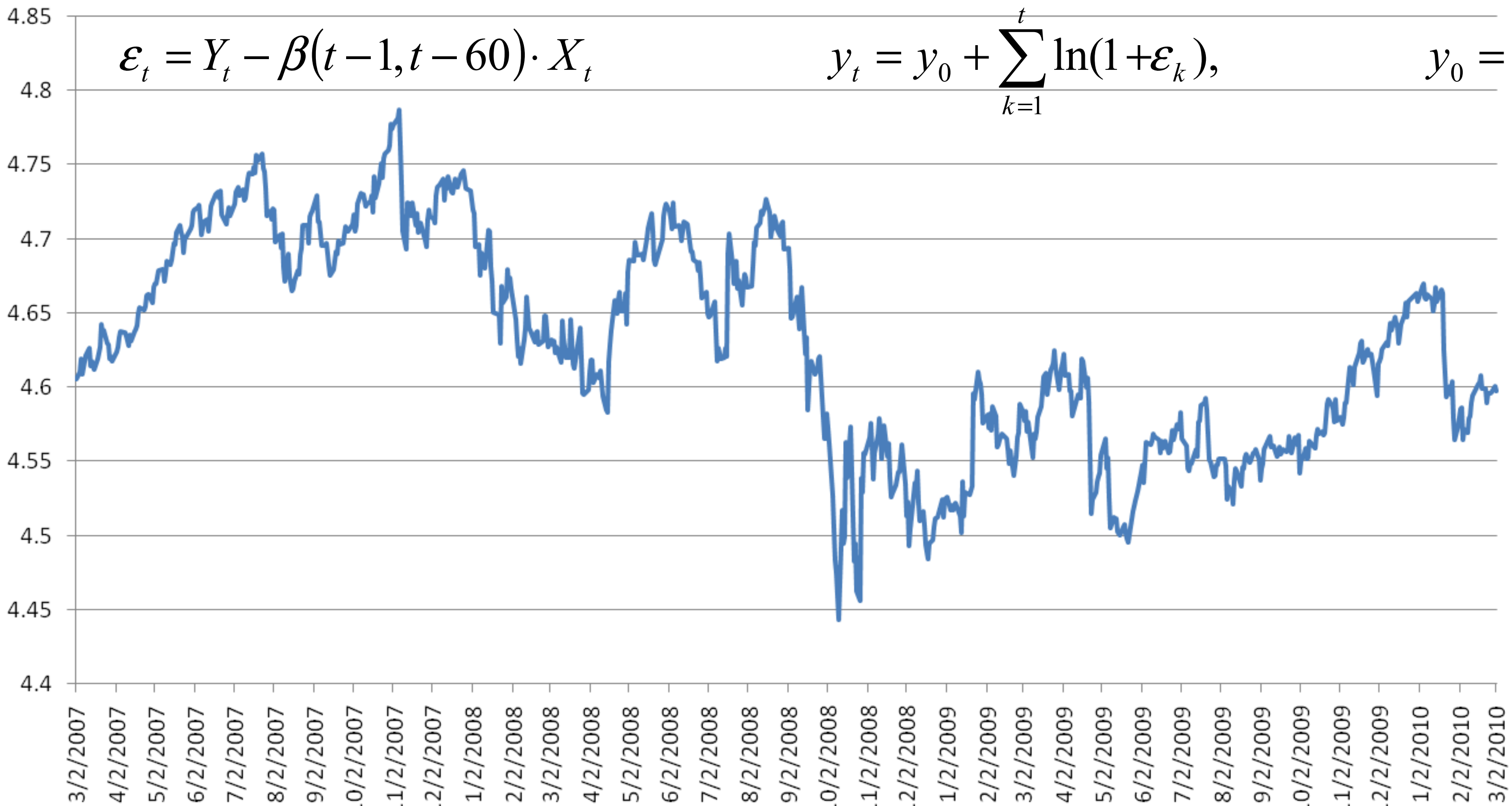
ADF Critical Values: Reject delta=0 if DF <	
1% level	-3.970385
5% level	-3.415895
10% level	-3.130187

EBAY vs. XLK residuals

$Y_t =$ daily return of EBAY

$X_t =$ daily return of XLK

$$\varepsilon_t = Y_t - \beta(t-1, t-60) \cdot X_t \qquad y_t = y_0 + \sum_{k=1}^{t} \ln(1+\varepsilon_k), \qquad y_0 = 100$$



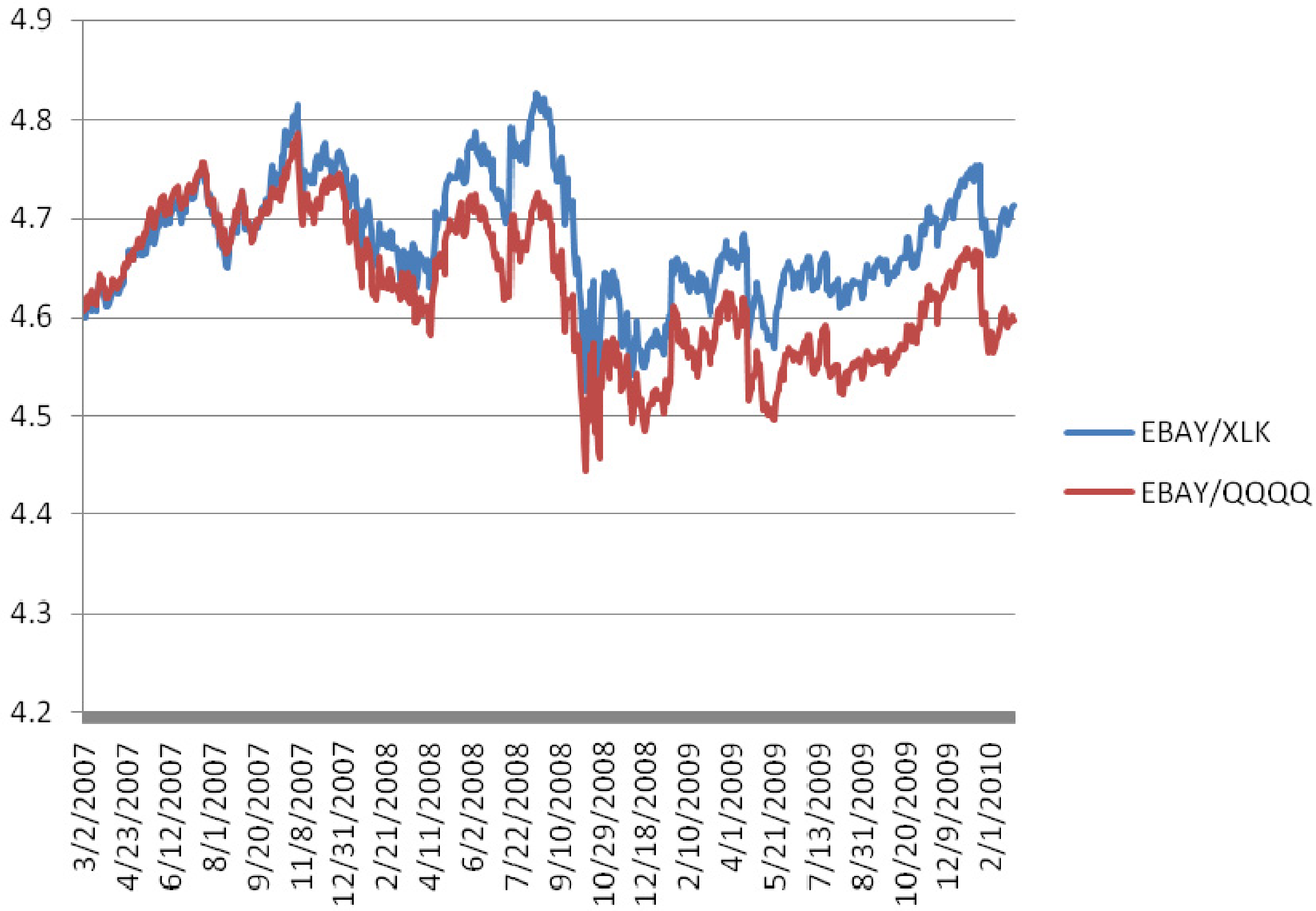
Augmented DF test for EBAY/XLK

Best lag fit: 9

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

Cannot reject UR @ 90% level

EBAY vs. QQQQ residuals



ADF for EBAY/QQQQ

Null Hypothesis: tseries has a unit root

Exogenous: Constant and linear Trend

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

Variable	Coefficient	Std. Error	t-Statistic	Prob
tseries(-1)	-0.023280	0.008338	-2.791940	0.005374
D(tseries(-1))	-0.078624	0.036419	-2.158873	0.031179
D(tseries(-2))	0.019488	0.036533	0.533428	0.593897
D(tseries(-3))	0.030306	0.036525	0.829726	0.406960
D(tseries(-4))	-0.114959	0.036359	-3.161785	0.001632
C	0.109870	0.039251	2.799187	0.005256
@trend	-0.000002	0.000002	-0.918302	0.358759

ARMA(p,q) process

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

Combines autorregressive models with moving average models

Simple linear time-series model

Fitting to an ARMA(1,1)

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

	Coefficient	Std. Error	t-Statistic	Prob.
c	4.627335411	0	148.9024	0
AR(1)	0.986154258	0	159.9401	0
MA(1)	-0.110605985	0	-2.998961	0.002798377

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

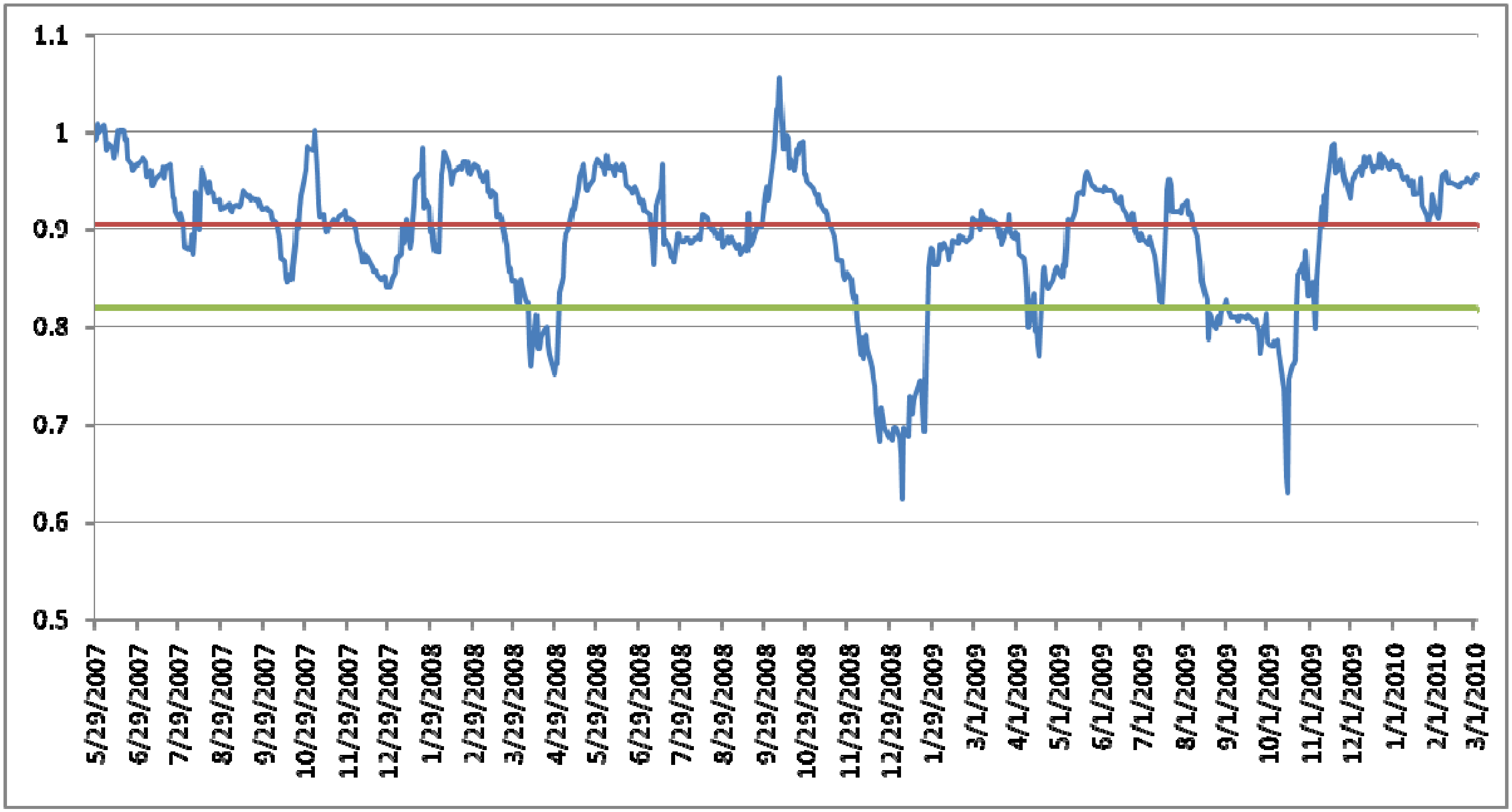
Included observations: 755

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

	Coefficient	Std. Error	t-Statistic	Prob.
c	4.627528	0.03	168.4630632	0
AR(1)	0.98229241	0.01	143.6624447	0

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:	1% level		-4.027516	
	5% level		-3.443485	
	10% level		-3.146482	
Variable	Coefficient	Std. Error	t-Statistic	Prob
tseries(-1)	-0.113728	0.043856	-2.593218	0.010671
D(tseries(-1))	-0.111532	0.090621	-1.230747	0.220785
D(tseries(-2))	0.162647	0.087448	1.859935	0.065303
D(tseries(-3))	0.040018	0.088750	0.450911	0.652854
D(tseries(-4))	-0.267631	0.085738	-3.121501	0.002247
D(tseries(-5))	0.076574	0.086639	0.883828	0.378528
D(tseries(-6))	-0.139433	0.085911	-1.623007	0.107169
D(tseries(-7))	-0.242743	0.082689	-2.935598	0.003980
D(tseries(-8))	0.090026	0.085786	1.049428	0.296056
D(tseries(-9))	0.189077	0.084225	2.244910	0.026575
D(tseries(-10))	-0.106441	0.084083	-1.265896	0.207962
C	0.511426	0.199365	2.565270	0.011521
@trend	0.000090	0.000044	2.058864	0.041636

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

	Coefficient	Std. Error	t-Statistic	Prob.
c	4.55489463	0.013841	329.0952386	0
AR(1)	0.88029659	0.035204	25.00582423	-2.2E-16

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