DEPARTMENT OF ECONOMICS
UNIVERSITY OF COPENHAGEN

Econometrics 2, Class 1

Problem Set #12
December 5, 2005

DEPARTMENT OF ECONOMICS
UNIVERSITY OF COPENHAGEN

From last time: PS#11
(7) Unrestricted error correction model

$$\Delta C_t = \alpha_0 + \alpha_1 \Delta C_{t-1} + \alpha_2 \Delta Y_t + \alpha_3 \Delta Y_{t-1} + \alpha_4 \Delta W_t + \alpha_5 \Delta W_{t-1} + \alpha_6 C_{t-1} + \alpha_7 Y_{t-1} + \alpha_8 W_{t-1} + \alpha_9 \text{ARBLOS}_t + \alpha_{10} Dum754_t + \alpha_{11} Dum773_t + \alpha_{12} Dum774_t + \epsilon_t.$$

EQ(18) Modelling DC by OLS (using ConsumptionData.in7)
The estimation sample is: 1973 (3) to 2003 (2)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
DC_1	-0.162208	0.07952	-2.04	0.044	0.0374
Constant	0.0344309	0.08075	0.426	0.671	0.0017
DY	0.0741451	0.05829	1.27	0.206	0.0149
DY_1	-0.0579263	0.05815	-0.996	0.321	0.0092
DW	0.321061	0.1044	3.07	0.003	0.0812
DW_1	0.0523076	0.1105	0.473	0.637	0.0021
C_1	-0.240254	0.06076	-3.95	0.000	0.1275
Y_1	0.111539	0.03897	2.86	0.005	0.0711
W_1	0.0902570	0.04038	2.24	0.027	0.0446
ARBLOS	-5.17642	1.356	-3.82	0.000	0.1198
DUM754	0.0631710	0.01467	4.31	0.000	0.1478
DUM773	0.0476979	0.01329	3.59	0.001	0.1074
DUM774	-0.0564696	0.01320	-4.28	0.000	0.1460

sigma	0.0123577	RSS	0.0163403029
R^2	0.610016	F(12,107) =	13.95 [0.000]**
log-likelihood	363.824	DW	1.9
no. of observations	120	no. of parameters	13
mean(DC)	0.00310834	var(DC)	0.000349166

HOMEWORK!

- Derive the long-run solution and compare with the ADL model.
- Perform PcGive test for no-cointegration.
- E-mail me if you have difficulties with this.
- I will expect someone to present their answer next week.

Estimating long-run solution
Repetition from lectures

The estimator of β_2 from a static regression is super-consistent...but
(1) $\widehat{\beta}_2$ is often biased (due to ignored dynamics).
(2) Hypotheses on β_2 cannot be tested.

An alternative estimator is based on an unrestricted ADL model, e.g.

$$X_{1t} = \delta + \theta_1 X_{1t-1} + \theta_2 X_{1t-2} + \phi_0 X_{2t} + \phi_1 X_{2t-1} + \phi_2 X_{2t-2} + \epsilon_t,$$

where ϵ_t is IID. This is equivalent to an error correction model:

$$\Delta X_{1t} = \delta + \lambda_1 \Delta X_{1t-1} + \kappa_0 \Delta X_{2t} + \kappa_1 \Delta X_{2t-1} + \gamma_1 X_{1t-1} + \gamma_2 X_{2t-1} + \epsilon_t.$$

An estimate of β_2 can be found from the long-run solutions:

$$\widehat{\beta}_2 = \frac{-\widehat{\gamma}_2}{\widehat{\gamma}_1} = \frac{\widehat{\phi}_0 + \widehat{\phi}_1 + \widehat{\phi}_2}{1 - \widehat{\theta}_1 - \widehat{\theta}_2}.$$

Estimating long-run solution

Long-run relationship: $C_t = \beta_0 + \beta_1 Y_t + \beta_2 W_t + u_t,$

ECM formulation:

$$\begin{aligned}\Delta C_t &= \alpha_0 + \alpha_1 \Delta C_{t-1} + \alpha_2 \Delta Y_t + \alpha_3 \Delta Y_{t-1} + \alpha_4 \Delta W_t + \alpha_5 \Delta W_{t-1} \\ &+ \alpha_6 C_{t-1} + \alpha_7 Y_{t-1} + \alpha_8 W_{t-1} + \alpha_9 \text{ARBLOS}_t \\ &+ \alpha_{10} Dum754_t + \alpha_{11} Dum773_t + \alpha_{12} Dum774_t + \epsilon_t.\end{aligned}$$

Use formula on last slide to obtain:

$$\beta_1 = -\frac{\alpha_7}{\alpha_6} = -\frac{0.111539}{-0.240254} = 0.464$$

$$\beta_2 = -\frac{\alpha_8}{\alpha_6} = -\frac{0.0902570}{-0.240254} = 0.376$$

Of course, these are the same results we derived using the ADL model!

DEPARTMENT OF ECONOMICS
UNIVERSITY OF COPENHAGEN

PcGive test for no-cointegration (Test – Dynamic Analysis – Lag structure analysis)

```
Tests on the significance of each variable
Variable      F-test          Value [  Prob]     Unit-root t-test
DC            F(1,107) =    4.1607 [0.0438]*          -14.615**
Constant      F(1,107) =   0.18181 [0.6707]
DY            F(2,107) =    2.0288 [0.1365]            0.1688
DW            F(2,107) =    4.8867 [0.0093]**          2.4941
C             F(1,107) =    15.633 [0.0001]**         -3.9539
Y             F(1,107) =    8.1908 [0.0051]**          2.862
W             F(1,107) =    4.9965 [0.0275]*           2.2353
ARBLOS        F(1,107) =    14.565 [0.0002]**         -3.8164
DUM754        F(1,107) =    18.555 [0.0000]**          4.3076
DUM773        F(1,107) =    12.878 [0.0005]**          3.5886
DUM774        F(1,107) =    18.299 [0.0000]**         -4.2777
```

N.B. Same as last time! We thus reject the null of no-cointegration.

(B) PcGive test for no-cointegration

Number of variables in X_t (p)	Constant in (25)			Constant and trend in (25)		
	1%	5%	10%	1%	5%	10%
2	−3.79	−3.21	−2.91	−4.25	−3.69	−3.39
3	−4.09	−3.51	−3.19	−4.50	−3.93	−3.62
4	−4.36	−3.76	−3.44	−4.72	−4.14	−3.83
5	−4.59	−3.99	−3.66	−4.93	−4.34	−4.03

Table 1: Asymptotic critical values for tests of no-cointegration. Reproduced from Davidson and MacKinnon (1993).

DEPARTMENT OF ECONOMICS
UNIVERSITY OF COPENHAGEN

#12.1 The ARCH Model

We will go through this old exam question on the blackboard.

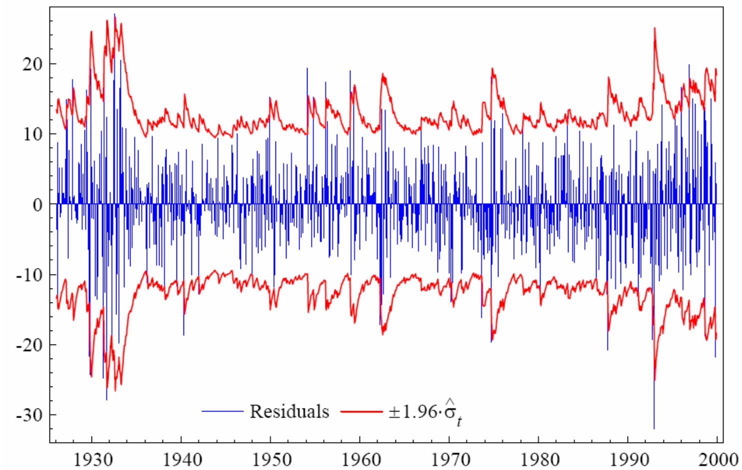


Figure 12.1: Estimated residuals, $\hat{\epsilon}_t$, and the conditional standard deviation.

DEPARTMENT OF ECONOMICS
UNIVERSITY OF COPENHAGEN

#12.2 Danish KFX Stock Market Index

In this exercise we consider the Danish KFX stock market index recorded weekly for the period week 4, 1993 to week 18, 2004. Let KFX be the stock market index, and define the log return

$$y_t = \log(\mathsf{KFX}_t) - \log(\mathsf{KFX}_{t-1}).$$

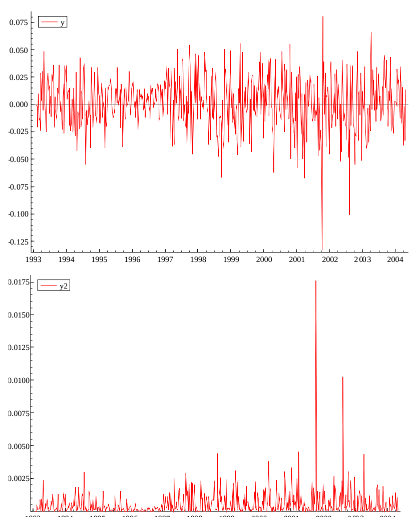
We want to estimate a GARCH model for y_t.

DEPARTMENT OF ECONOMICS
UNIVERSITY OF COPENHAGEN

(1) Graph of stock index

DEPARTMENT OF ECONOMICS
UNIVERSITY OF COPENHAGEN

Graphs of log returns and squared log returns



There do seem to be periods of increased volatility/variance. This is symptomatic of ARCH effects.

(2) Autoregressive model for *y*

We try an AR(4) model:

$$y_t = \delta + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \theta_3 Y_{t-3} + \theta_4 Y_{t-4} + \varepsilon_t$$

```
EQ( 1) Modelling y by OLS (using kfx.in7)
       The estimation sample is: 1993 (9) to 2004 (18)

                  Coefficient  Std.Error  t-value  t-prob Part.R^2
y_1                 0.0478563    0.04141     1.16   0.248   0.0023
y_2                 0.0565597    0.04138     1.37   0.172   0.0032
y_3                 0.0650485    0.04146     1.57   0.117   0.0042
y_4                -0.0961051    0.04149    -2.32   0.021   0.0092
Constant           0.00186557  0.0009954     1.87   0.061   0.0061

sigma               0.0237355  RSS             0.32506809
R^2                  0.018069  F(4,577) =     2.654 [0.032]*
log-likelihood        1353.82  DW                    2.01
no. of observations       582  no. of parameters        5
mean(y)            0.00200738  var(y)         0.000568814
```

The lags are generally insignificant (good theoretically!)...

Autoregressive model for *y* (cont.)

... so we remove one lag at a time and end up with this model:

$$y_t = \delta + \varepsilon_t$$

```
EQ( 2) Modelling y by OLS (using kfx.in7)
       The estimation sample is: 1993 (9) to 2004 (18)

                  Coefficient  Std.Error  t-value  t-prob Part.R^2
Constant           0.00200738  0.0009895     2.03   0.043   0.0070

sigma               0.0238703  RSS            0.331049823
R^2               2.88042e-033
log-likelihood        1348.52  DW                     1.9
no. of observations       582  no. of parameters        1
mean(y)            0.00200738  var(y)         0.000568814

AR 1-7 test:      F(7,574)  =   2.6477 [0.0106]*
ARCH 1-7 test:    F(7,567)  =   3.5047 [0.0011]**
Normality test:   Chi^2(2)  =   48.367 [0.0000]**
RESET test:       F(1,580)  =   4.4027 [0.0363]*
```

We should really have removed this, but try yourselves! (It is only borderline significant.)

The model is very poorly specified (not a surprise). We can however use it to test for ARCH effects in the residuals – *if there is no residual autocorrelation*.

Why is it important that we remove AR before testing for ARCH?

Autocorrelation: There is a relationship between the residuals and the lagged residuals.

This implies that there is also a relationship between the squared residuals and the lagged squared residuals.

ARCH: There is a relationship between the squared residuals and the lagged squared residuals.

So AR => ARCH
BUT
ARCH does ***not*** imply AR!

(3) Test – Test… - ARCH test (Breusch-Pagan test for heteroskedasticity)

To test for ARCH of order p consider the auxiliary regression model

$$\epsilon_t^2 = \beta_0 + \beta_1 \epsilon_{t-1}^2 + \beta_2 \epsilon_{t-2}^2 + ... + \beta_p \epsilon_{t-p}^2 + \eta_t.$$

Under the null of no ARCH,

$$H_0 : \beta_1 = \beta_2 = ... = \beta_p = 0.$$

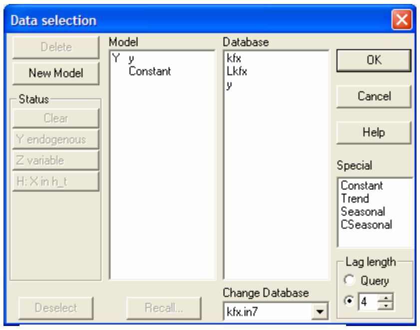
```
ARCH coefficients:
  Lag Coefficient  Std.Error
    1      0.16831    0.04197
    2     0.012206     0.0424
    3     0.058302    0.04241
    4    0.0063574     0.0425
    5    -0.023318    0.04246
    6     0.089259    0.04248
    7    -0.034566    0.04203
RSS = 0.000713625  sigma = 0.00112187

Testing for error ARCH from lags 1 to 7
ARCH 1-7 test:    F(7,567) =   3.5047 [0.0011]**
```

In the presence of ARCH, OLS is consistent but inefficient.

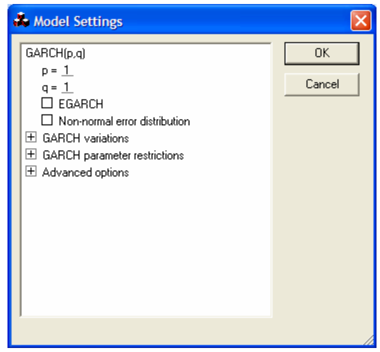
This is the same test presented in the previous slide. The null hypothesis is no ARCH, which we reject. (N.B. PcGive reports F-statistics, not chi-sq)

(4) Package – Volatility Models
Model – Formulate...

Specify the conditional mean:



Specify the conditional variance:



$$\sigma_t^2 = \varpi + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

We choose the GARCH(1,1) model for the conditional variance. This nearly always fits, although it is quite restrictive. (Implies an ARMA(1,1) model for the squared innovation.)

Results (using ML and assuming normality)

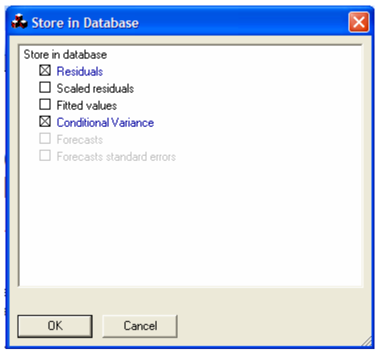
```
VOL( 1) Modelling y by restricted GARCH(1,1) (kfx.in7)
        The estimation sample is: 1993 (5) to 2004 (18)

                  Coefficient  Std.Error  robust-SE  t-value  t-prob
Constant      X    0.00302607  0.0008463  0.0008175     3.70   0.000
alpha_0       H  3.25528e-006 3.285e-006 2.821e-006     1.15   0.249
alpha_1       H     0.0733465    0.02065    0.02418     3.03   0.003
beta_1        H      0.924945    0.02083    0.02252     41.1   0.000

log-likelihood     1382.50548  HMSE                  2.81717
mean(h_t)         0.000600046  var(h_t)         9.81728e-008
no. of observations       586  no. of parameters           4
AIC.T             -2757.01097  AIC               -4.70479687
mean(y)            0.00195163  var(y)            0.000565861
alpha(1)+beta(1)     0.998291  alpha_i+beta_i>=0, alpha(1)+beta(1)<1
```

X indicates a coefficient from the conditional mean (where we only have a constant).

H indicates a coefficient from the conditional variance (constant, alpha and beta).

(Note: We (almost) have a non-stationary ARMA process for the squared residuals. Persistence in volatility is very high!)

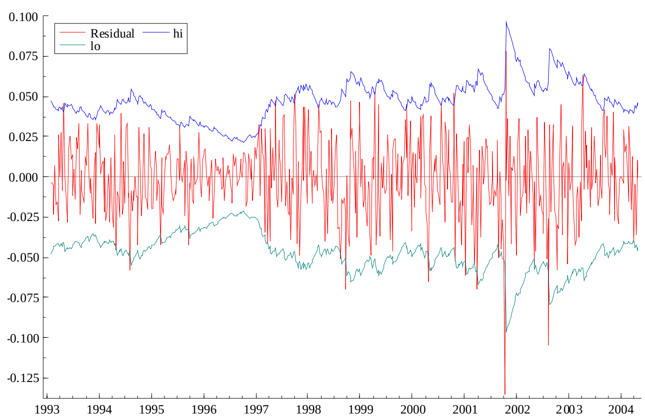
(5) Compare conditional standard deviation with the residuals



- Save the residuals and the conditional variance.
- Define the conditional standard deviation.
- Error bands for the residuals can be constructed as $\pm 2\sigma_t$
- Save these as e.g. hi and lo.
- You can now draw the graph on the following slide.

DEPARTMENT OF ECONOMICS
UNIVERSITY OF COPENHAGEN

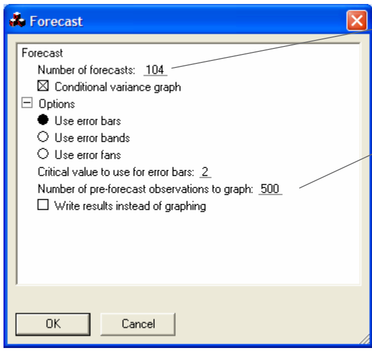
The pretty picture...



Note the typical ARCH pattern.

Large shocks are followed by periods of increased volatility.

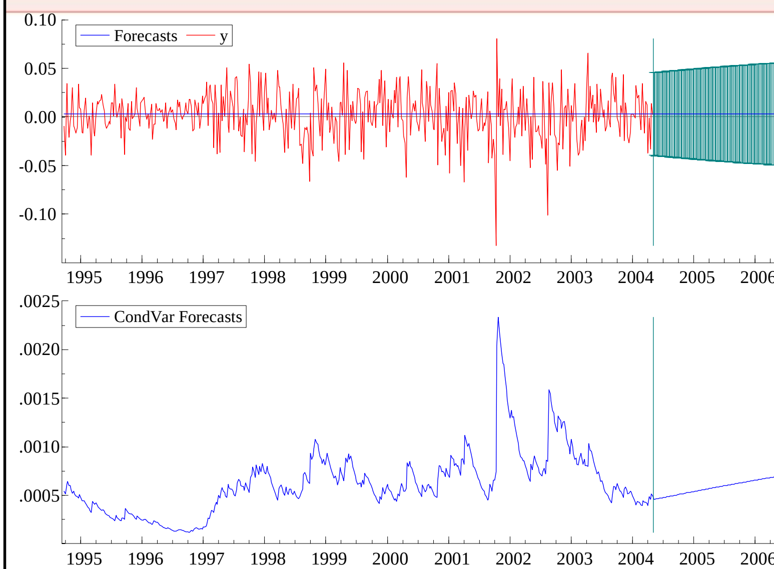
(6) Make a 52 period out of sample forecast



I have chosen a larger number than suggested.

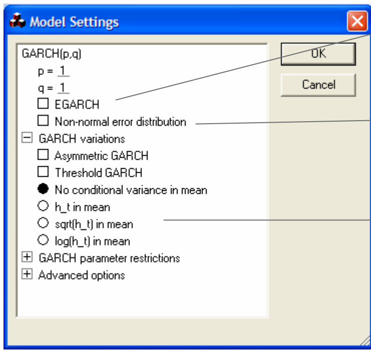
Remember to increase this!

The forecast results



With no shocks, the conditional variance will gradually adjust back to the unconditional variance. (History means less and less.)

(7) Try other specifications for the conditional variance.



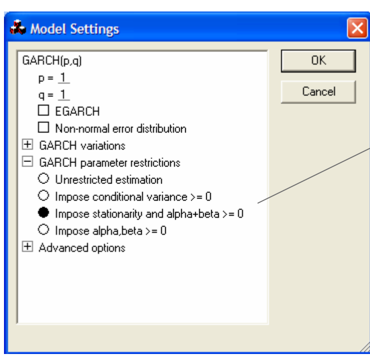
Allows shocks to have asymmetric effects. (e.g. negative shocks lead to more volatility)

You could try a non-normal distribution for the errors (t-distribution).

h_t is the conditional variance. These options allow volatility to impact on the conditional mean. (e.g. higher risk means a higher risk premium is necessary)

DEPARTMENT OF ECONOMICS
UNIVERSITY OF COPENHAGEN

Other specifications (cont.)



You can also impose stationarity, positive conditional variance etc.