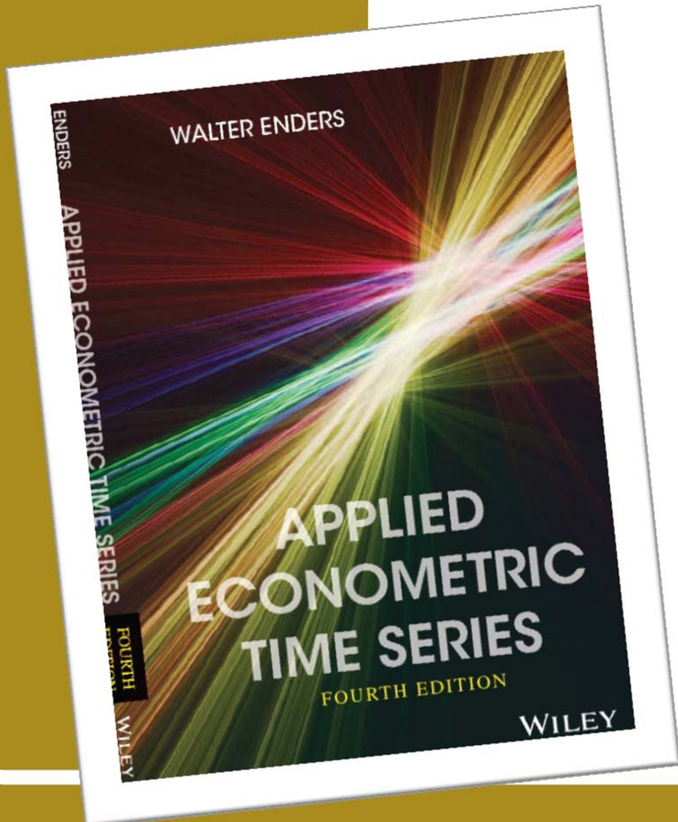


APPLIED ECONOMETRIC TIME SERIES 4TH ED.

Chapter 3: *Modeling Volatility*

WALTER ENDERS, UNIVERSITY OF ALABAMA

Section 1

ECONOMIC TIME SERIES: THE STYLIZED FACTS

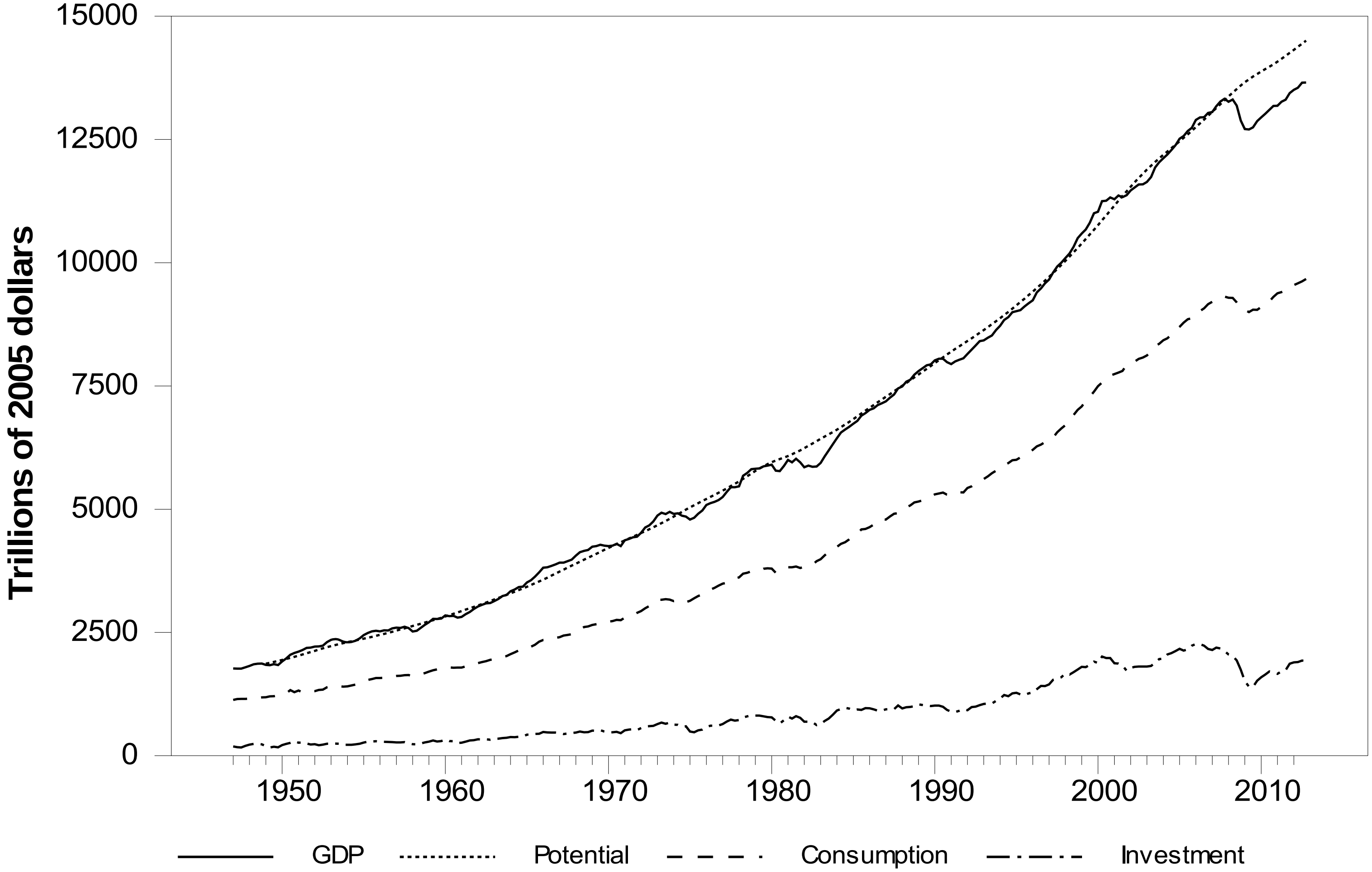


Figure 3.1 Real GDP, Consumption and Investment

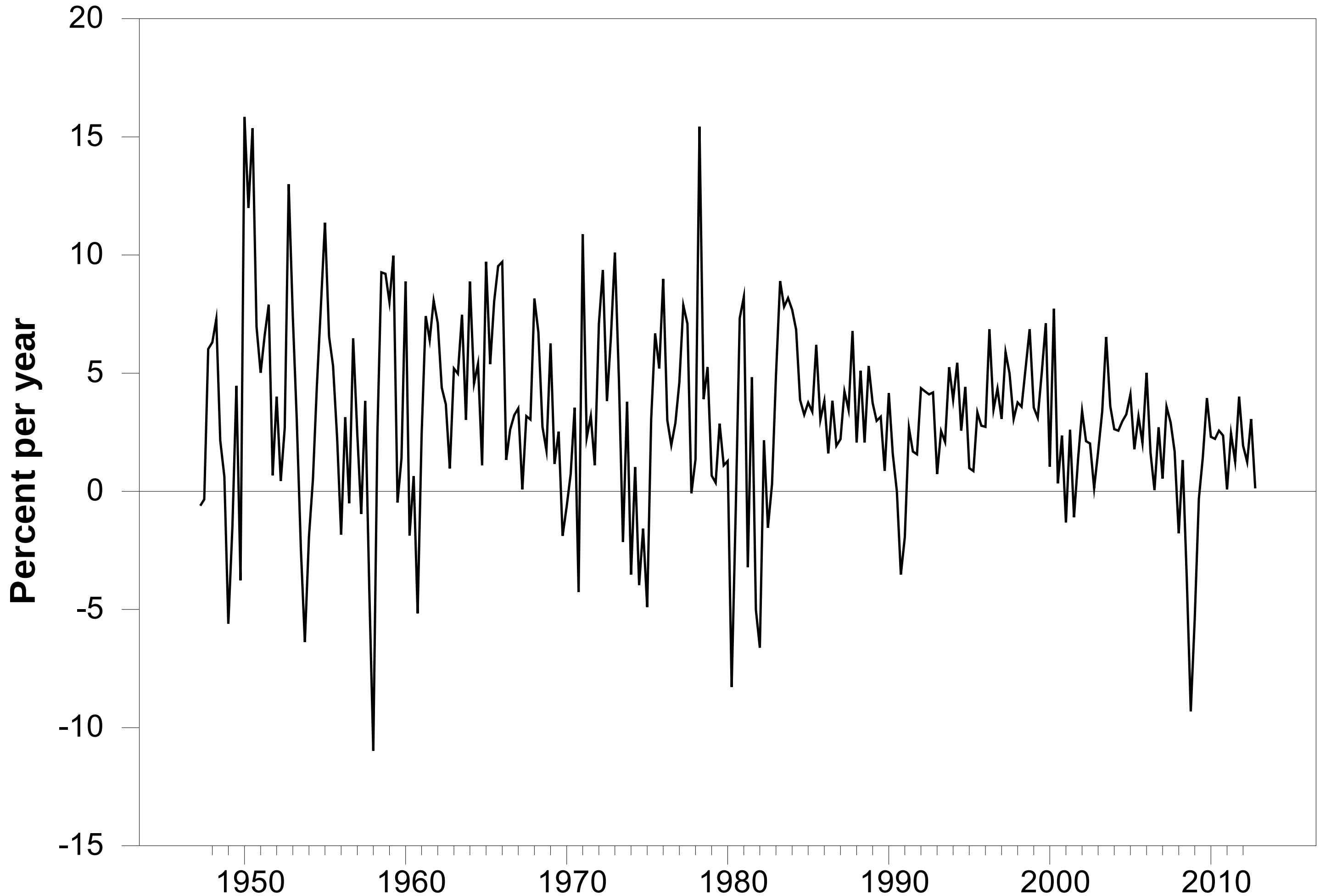


Figure 3.2 Annualized Growth Rate of Real GDP

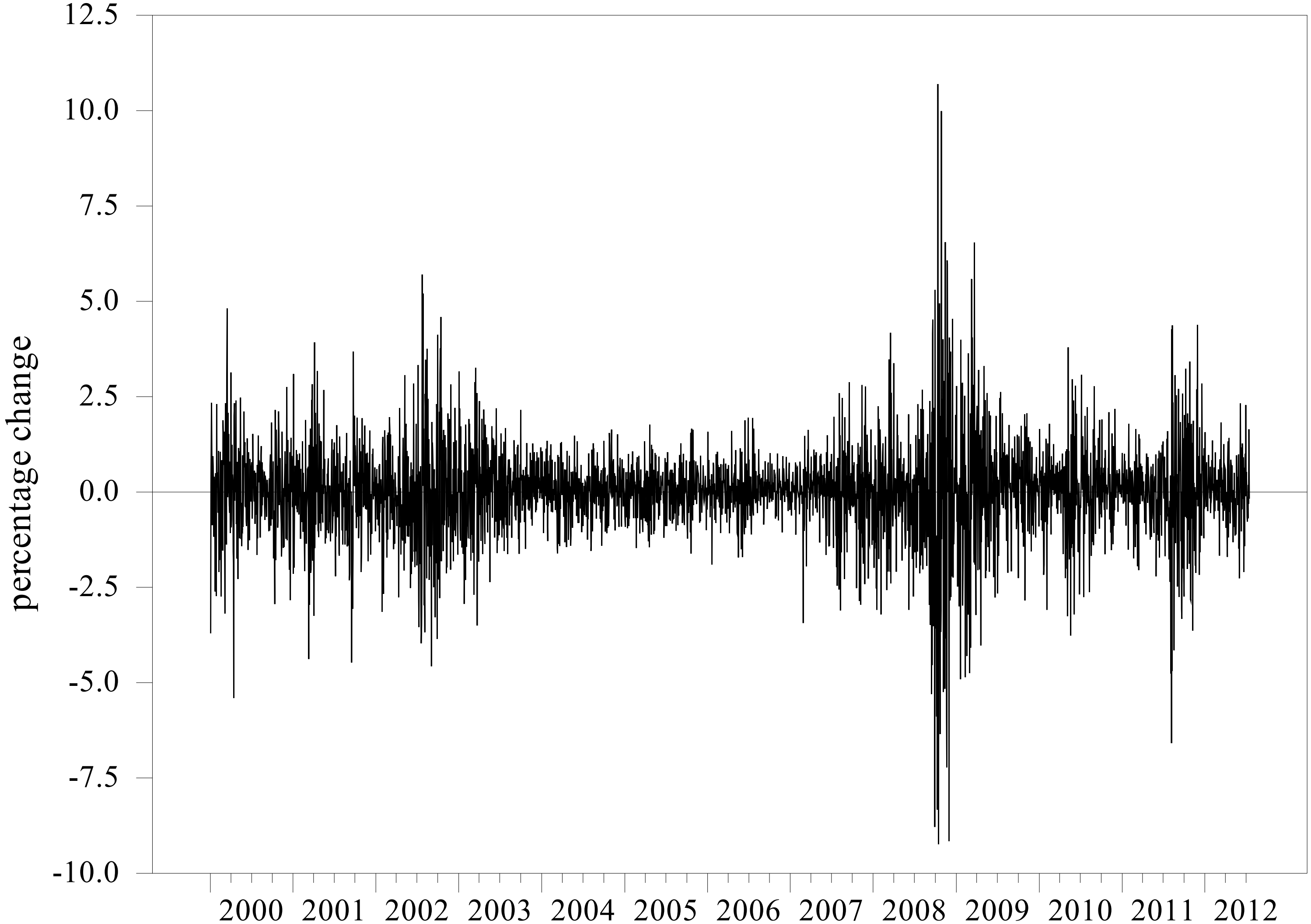


Figure 3.3: Percentage Change in the NYSE US 100: (Jan 4, 2000 - July 16, 2012)

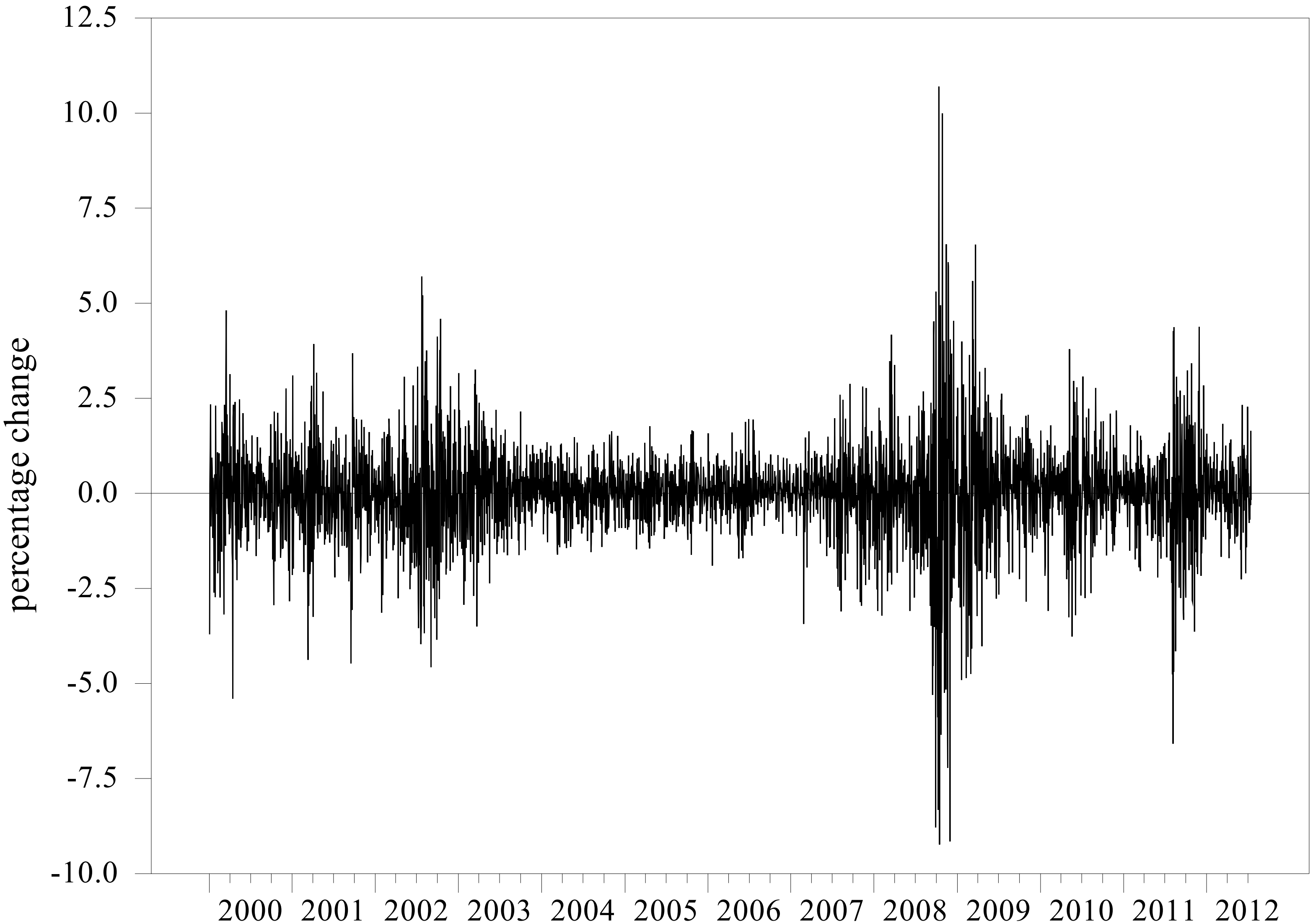


Figure 3.3: Percentage Change in the NYSE US 100: (Jan 4, 2000 - July 16, 2012)

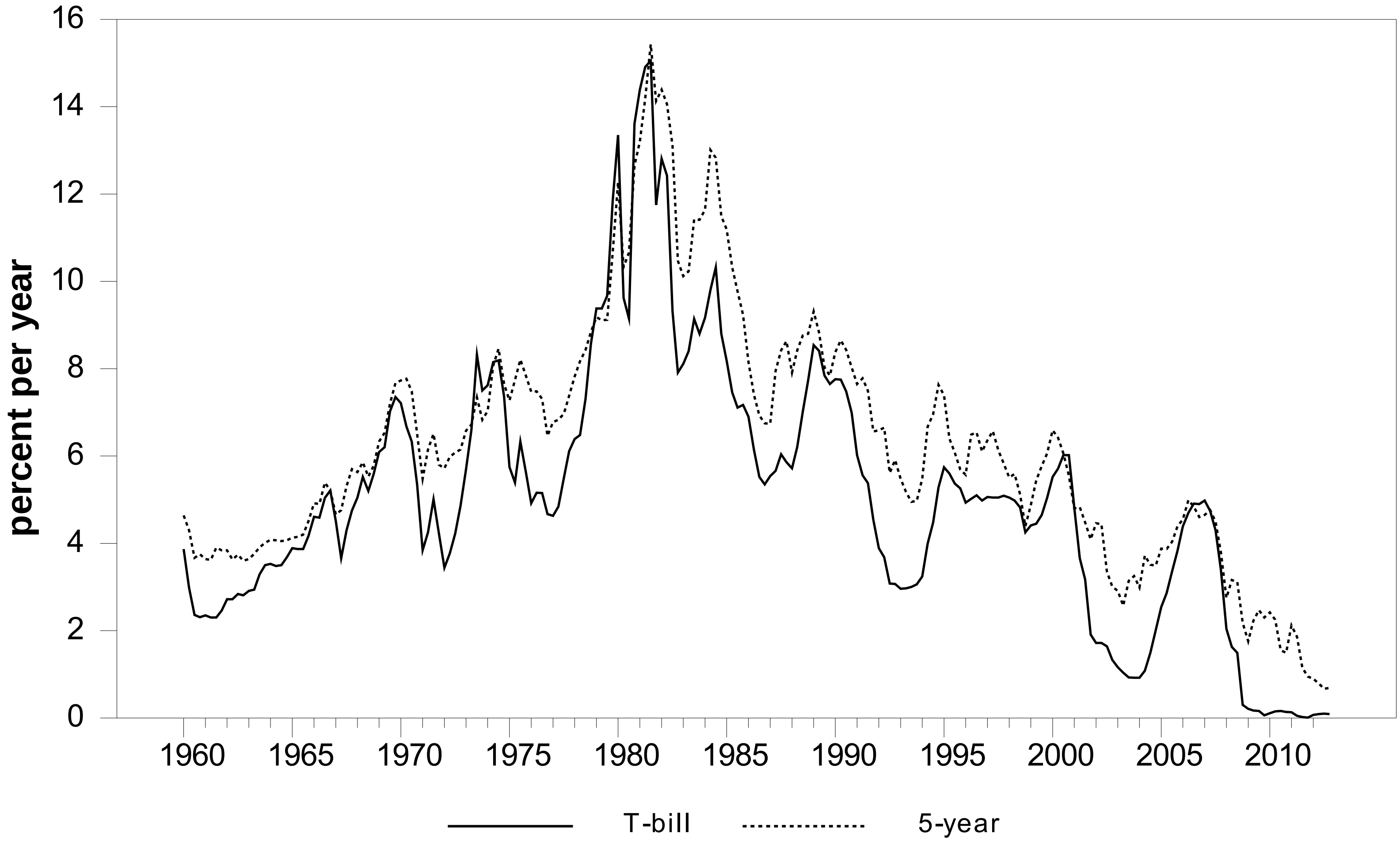


Figure 3.4 Short- and Long-Term Interest Rates

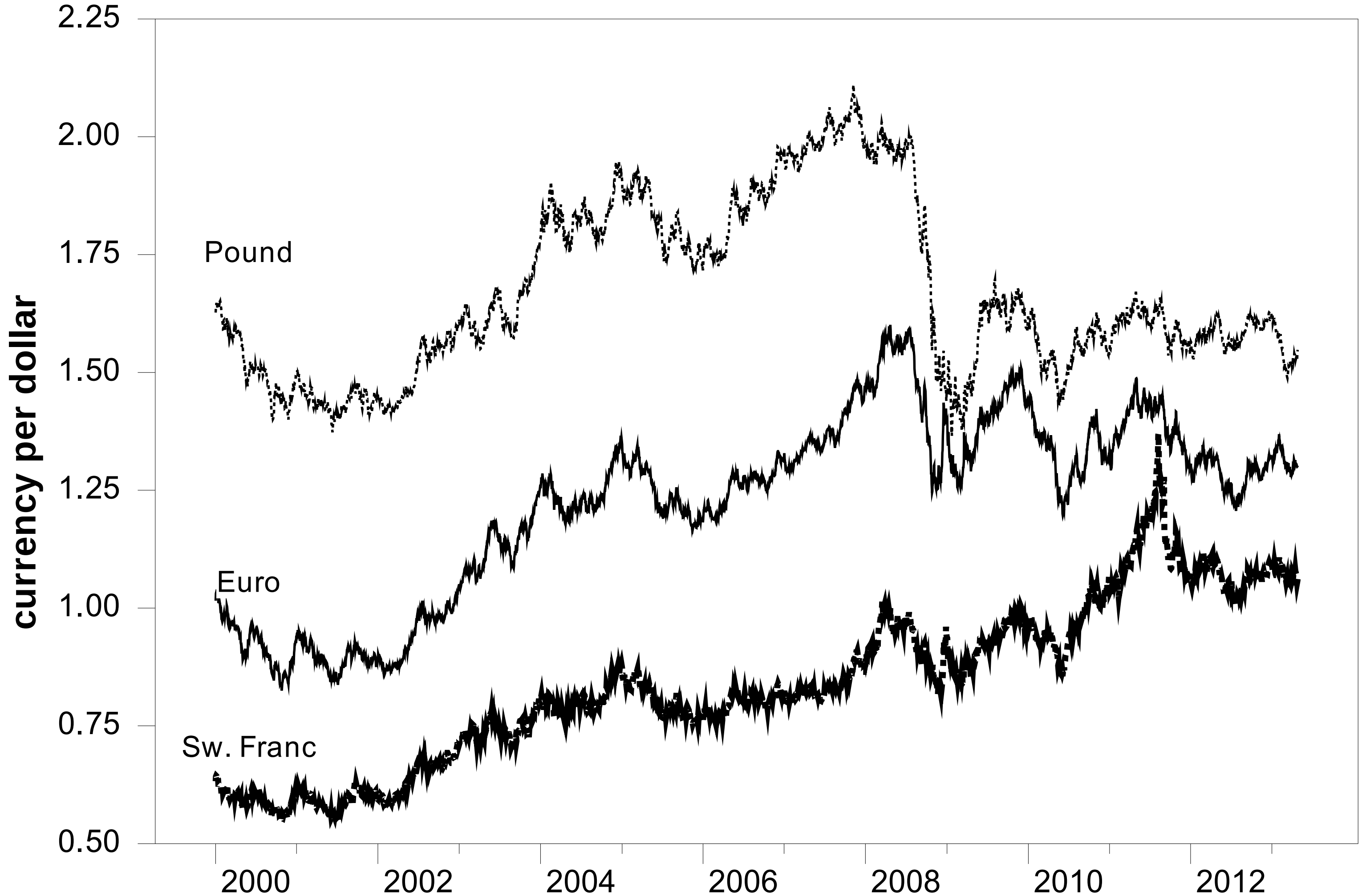


Figure 3.5: Daily Exchange Rates (Jan 3, 2000 - April 4, 2013)

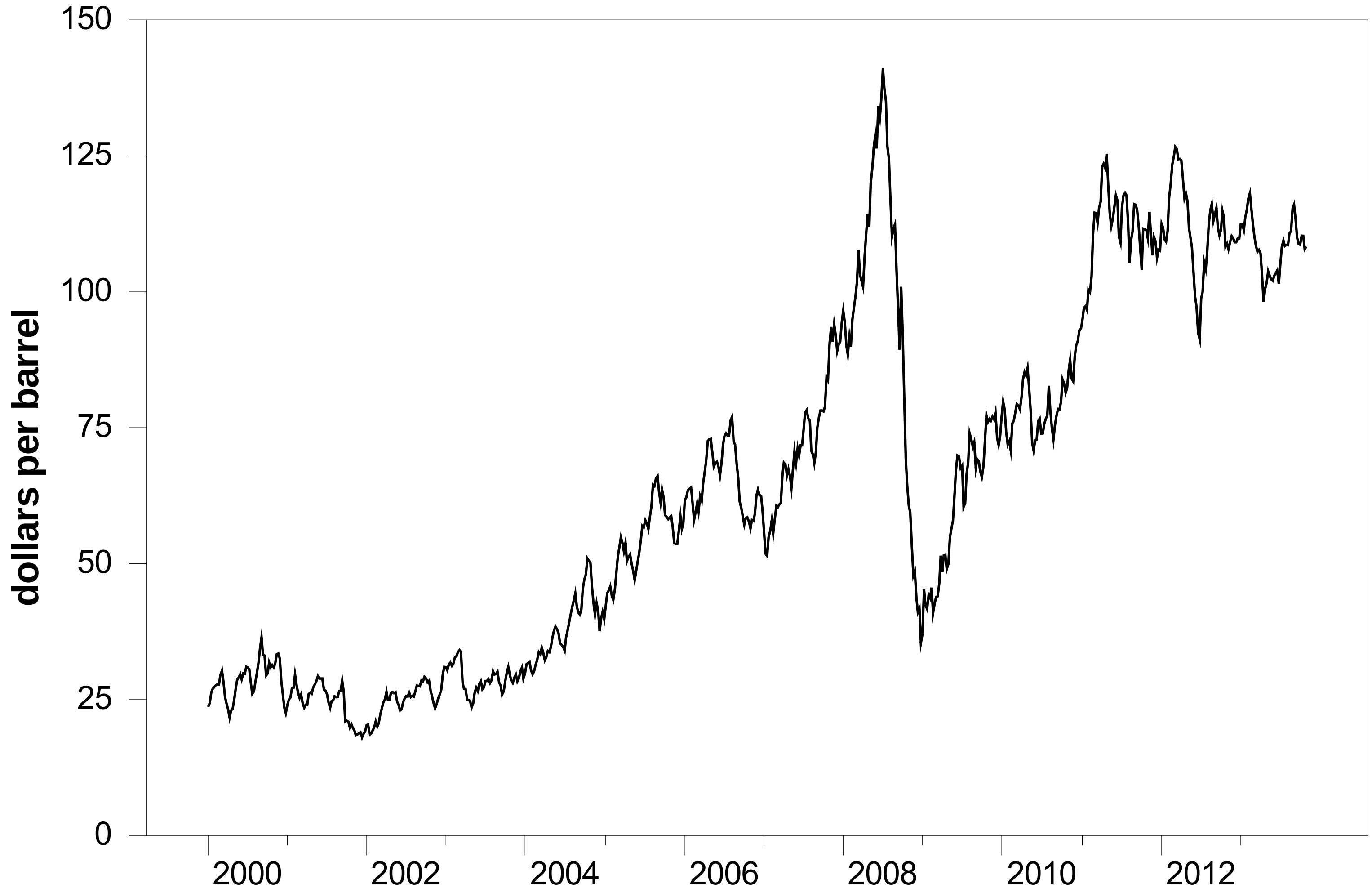


Figure 3.6: Weekly Values of the Spot Price of Oil: (May 15, 1987 - Nov 1, 2013)

ARCH Processes

The GARCH Model

2. ARCH AND GARCH PROCESSES

Other Methods

- Let ε_t = demeanded daily return. One method is to use 30-day moving average $\sum_1^{30} {\varepsilon_t}^2/30$

- Implicit volatility

- Logs can stabilize volatility

One simple strategy is to model the conditional variance as an AR(q) process using *squares* of the estimated residuals

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \ldots + \alpha_q \hat{\varepsilon}_{t-q}^2 + v_t$$

In contrast to the moving average, here the weights need not equal1/30 (or 1/N).

The forecasts are:

$$E_t \hat{\varepsilon}_{t+1}^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_t^2 + \alpha_2 \hat{\varepsilon}_{t-1}^2 + \ldots + \alpha_q \hat{\varepsilon}_{t+1-q}^2$$

Properties of the Simple ARCH Model

$$\varepsilon_t = v_t \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2}$$

Since v_t and ε_{t-1} are independent:

$$E\varepsilon_t = E[\ v_t(\alpha_0 + \alpha_1\varepsilon_{t-1}^2\ )]^{1/2}\ ] = 0$$

$$E_{t-1}\varepsilon_t = E_{t-1}v_tE_{t-1}\ [\alpha_0 + \alpha_1\varepsilon_{t-1}^2\ ]^{1/2}\ ] = 0$$

$$\mathrm{E}\varepsilon_t\ \varepsilon_{t-i} = 0 \quad (\ i \neq 0)$$

$$E\varepsilon_t{}^2 = E[\ v_t{}^2(\alpha_0 + \alpha_1\varepsilon_{t-1}^2\ )] = \alpha_0 + \alpha_1 E(\varepsilon_{t-1})^2$$

$$= \alpha_0/(\ 1 - \alpha_1\ )$$

$$E_{t-1}\varepsilon_t{}^2 = E_{t-1}\ [\ v_t{}^2(\alpha_0 + \alpha_1(\varepsilon_{t-1}\ )^2\ )] = \alpha_0 + \alpha_1(\varepsilon_{t-1})^2$$

ARCH Interactions with the Mean

Consider: $y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$

$$\text{Var}(y_t \mid y_{t-1}, y_{t-2}, \ldots) = E_{t-1}(y_t - a_0 - a_1 y_{t-1})^2$$

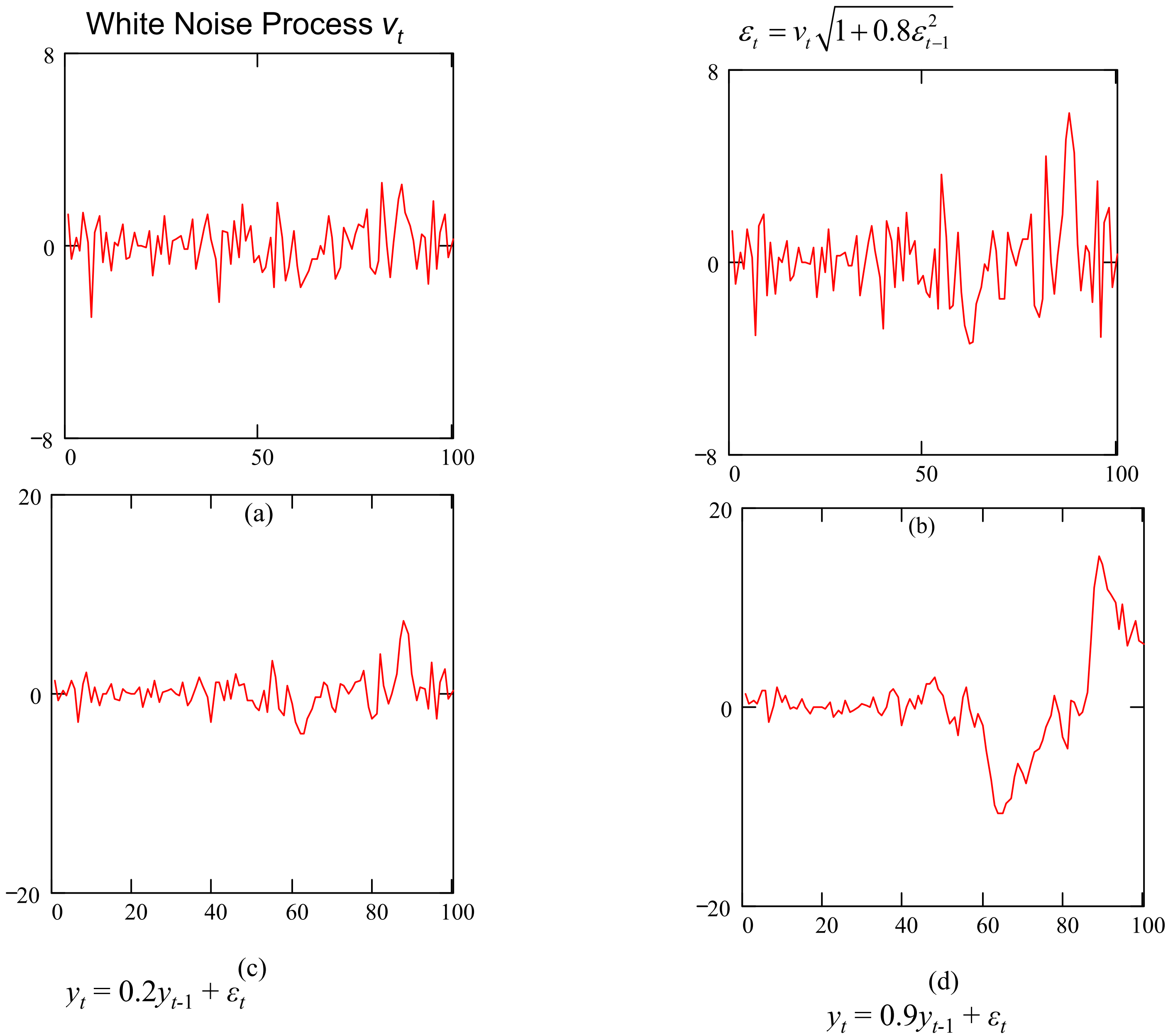
$$= E_{t-1}(\varepsilon_t)^2 = \alpha_0 + \alpha_1(\varepsilon_{t-1})^2$$

Unconditional Variance:

Since: $$y_t = \frac{a_0}{1 - a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i} \qquad \text{var}(y_t) = \sum_{i=0}^{\infty} a_1^{2i} \text{var}(\varepsilon_{t-i})$$

$$= \left(\frac{\alpha_0}{1 - \alpha_1}\right)\left(\frac{1}{1 - a_1^2}\right)$$

Figure 3.7: Simulated ARCH Processes



Other Processes

ARCH(q)

$$\varepsilon_t = v_t \sqrt{\alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2}$$

GARCH(p, q)

$$\varepsilon_t = v_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i}$$

The benefits of the GARCH model should be clear; a high-order ARCH model may have a more parsimonious GARCH representation that is much easier to identify and estimate. This is particularly true since all coefficients must be positive.

Testing For ARCH

- **Step 1**: Estimate the $\{y_t\}$ sequence using the "best fitting" ARMA model (or regression model) and obtain the squares of the fitted errors . Consider the regression equation:

$$\varepsilon_t^2 = a_0 + a_1\varepsilon_{t-1}^2 + a_2\varepsilon_{t-2}^2 + \ldots$$

If there are no ARCH effects $a_1 = a_2 = \ldots = 0$

- All the coefficients should be statistically significant
- No simple way to distinguish between various ARCH and GARCH models

Testing for ARCH II

- Examine the ACF of the squared residuals:
  - Calculate and plot the sample autocorrelations of the squared residuals
  - Ljung–Box *Q*-statistics can be used to test for groups of significant coefficients.

$$Q = T(T+2)\sum_{i=1}^{n}\rho_i^2/(T-i)$$

Q has an asymptotic χ^2 distribution with n degrees of freedom

Engle's Model of U.K. Inflation

- Let π = inflation and r = real wage

$$\pi_t = 0.0257 + 0.334\pi_{t-1} + 0.408\pi_{t-4} - 0.404\pi_{t-5} + 0.0559r_{t-1} + \varepsilon_t$$

the variance of ε_t is $h_t = 8.9 \text{ x } 10^{-5}$

$$\pi_t = 0.0328 + 0.162\pi_{t-1} + 0.264\pi_{t-4} - 0.325\pi_{t-5} + 0.0707r_{t-1} + \varepsilon_t$$

$$h_t = 1.4 \text{ x } 10^{-5} + 0.955(0.4 + 0.3L + 0.2L^2 + 0.1L^3)(\varepsilon_{t-1})^2$$

$$(8.5 \text{ x } 10^{-6}) \quad (0.298)$$

A GARCH Model of Oil Prices

Volatility Moderation

A GARCH Model of the Spread

4. THREE EXAMPLES OF GARCH MODELS

A GARCH Model of Oil Prices

- Use OIL.XLS to create

$$p_t = 100.0*[\log(spot_t) - \log(spot_{t-1})].$$

The following MA model works well:

$$p_t = 0.127 + \varepsilon_t + 0.177\varepsilon_{t-1} + 0.095\varepsilon_{t-3}$$

The McLeod–Li (1983) test for ARCH errors using four lags:

$$\varepsilon_t^2 = a_0 + a_1\varepsilon_{t-1}^2 + a_2\varepsilon_{t-2}^2 + \ldots$$

The *F*-statistic for the null hypothesis that the coefficients α_1 through α_4 all equal zero is 26.42. With 4 numerator and 1372 denominator degrees of freedom, we reject the null hypothesis of no ARCH errors at any conventional significance level.

The GARCH(1,1) Model

$$p_t = 0.130 + \varepsilon_t + 0.225\varepsilon_{t-1}$$

$$h_t = 0.402 + 0.097\,(\varepsilon_{t-1})^2 + 0.881h_{t-1}$$

Volatility Moderation

Use the file RGDP.XLS to construct the growth rate of real U.S. GDP

$y_t = \log(\text{RGDP}_t/\text{RGDP}_{t-1})$.

A reasonable model is:

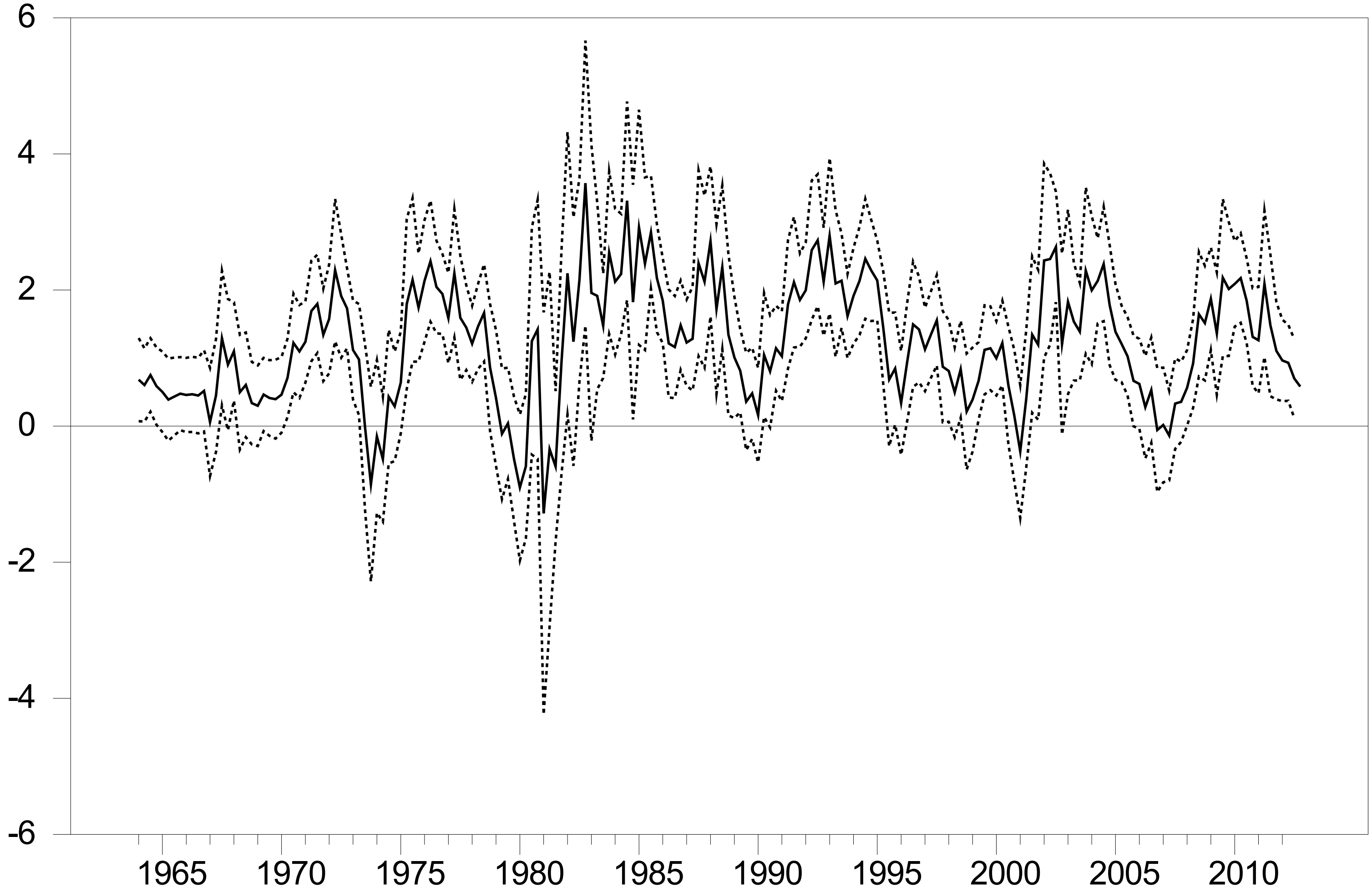
$$y_t = 0.005 + 0.371 y_{t-1} + \varepsilon_t$$
$$(6.80) \quad (6.44)$$

After checking for GARCH errors

$$y_t = 0.004 + 0.398 y_{t-1} + \varepsilon_t$$
$$(7.50) \quad (6.76)$$
$$h_t = 1.10\text{x}10^{-4} + 0.182 - 8.76\text{x}10^{-5} D_t$$
$$(7.87) \qquad (2.89) \qquad (-6.14)$$

The intercept of the variance equation was 1.10×10^{-4} prior to 1984Q1 and experienced a significant decline to 2.22×10^{-5} ($= 1.10 \times 10^{-4} - 8.76 \times 10^{-5}$) beginning in 1984$Q$1.

Figure 3.8: Forecasts of the Spread



_____ 1-Step Ahead Forecast - - - - - ± 2 Conditional Standard Deviations

Section 5

A GARCH MODEL OF RISK

Holt and Aradhyula (1990)

- The study examines the extent to which producers in the U.S. broiler (i.e., chicken) industry exhibit risk averse behavior.
  - The supply function for the U.S. broiler industry takes the form:

$$q_t = a_0 + a_1 p^e{}_t - a_2 h_t - a_3 pfeed_{t\text{-}1} + a_4 hatch_{t\text{-}1} + a_5 q_{t\text{-}4} + \varepsilon_{1t}$$

q_t = quantity of broiler production (in millions of pounds) in t;
$p^e{}_t = E_{t\text{-}1} p_t$ = expected real price of broilers at t
h_t = expected variance of the price of broilers in t
$pfeed_{t\text{-}1}$ = real price of broiler feed (in cents per pound) at t-1;
$hatch_{t\text{-}1}$ = hatch of broiler-type chicks in t-1;
ε_{1t} = supply shock in t;

and the length of the time period is one quarter.

- Note the negative effect of the conditional variance of price on broiler supply.
- The timing of the production process is such that feed and other production costs must be incurred before output is sold in the market.
- Producers must forecast the price that will prevail two months hence.
- The greater p_t^e, the greater the number of chicks that will be fed and brought to market.
- If price variability is very low, these forecasts can be held with confidence. Increased price variability decreases the accuracy of the forecasts and decreases broiler supply.
  - Risk-averse producers will opt to raise and market fewer broilers when the conditional volatility of price is high.

The Price equation

- $(1 - 0.511L - 0.129L^2 - 0.130L^3 - 0.138\mathrm{L}^4)p_t = 1.632 + \varepsilon_{2t}$
- $h_t = 1.353 + 0.162\varepsilon_{2t\text{-}1} + 0.591h_{t\text{-}1}$
- The paper assumes producers use these equations to form their price expectations. The supply equation:

$$q_t = 2.767p^e{}_t - 0.521E_{t\text{-}1}h_t - 4.325pfeed_{t\text{-}1}$$

$$+ 1.887hatch_{t\text{-}1} + 0.603p_{t\text{-}4} + \varepsilon_{1t}$$

Section 6: THE ARCH-M MODEL

- Engle, Lilien, and Robins let

$$y_t = \mu_t + \varepsilon_t$$

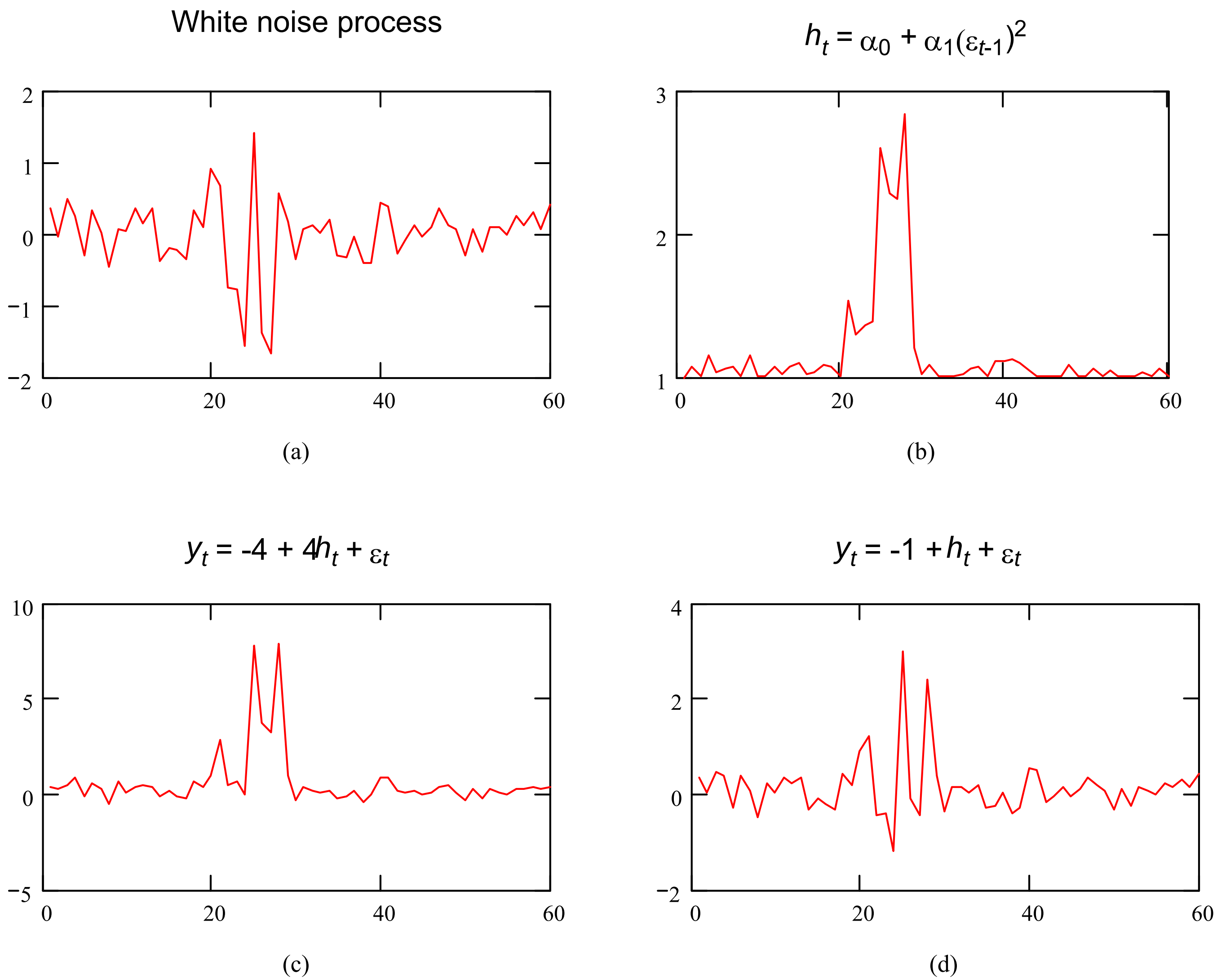
where y_t = excess return from holding a long-term asset relative to a one-period treasury bill and μ_t is a time-varying risk premium: $E_{t-1}y_t = \mu_t$

The risk premium is:

$$\mu_t = \beta + \delta h_t , \qquad \delta > 0$$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$$

Figure 3.9: Simulated ARCH-M Processes



Diagnostic Checks for Model Adequacy

Forecasting the Conditional Variance

7. ADDITIONAL PROPERTIES OF GARCH PROCESS

Forecasting with the GARCH(1, 1)

1. $E_T\varepsilon_{T+1}^2 = h_{T+1} = \alpha_0 + \alpha_1\varepsilon_T^2 + \beta_1 h_T$

The 1-step ahead forecast can be calculated directly.

2. Variance: $\varepsilon_t^2 = v_t^2(\alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1 h_{t-1})$

$$E\varepsilon_t^2 = Ev_t^2(\alpha_0 + \alpha_1 E\varepsilon_{t-1}^2 + \beta_1 Eh_{t-1})$$

$$Note: E\varepsilon_{t-1}^2 = E(E_{t-2}\varepsilon_{t-1}^2) = Eh_{t-1}$$

$$E\varepsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1)E\varepsilon_{t-1}^2 = \alpha_0/(1 - \alpha_1 - \beta_1)$$

3. Mult-Step Forecasts: Note that $E_t\varepsilon_{t+j}^2 = E_t h_{t+j}$

$$E_t h_{t+j} = \alpha_0 + \alpha_1 E_t\varepsilon_{t+j-1}^2 + \beta_1 E_t h_{t+j-1}$$

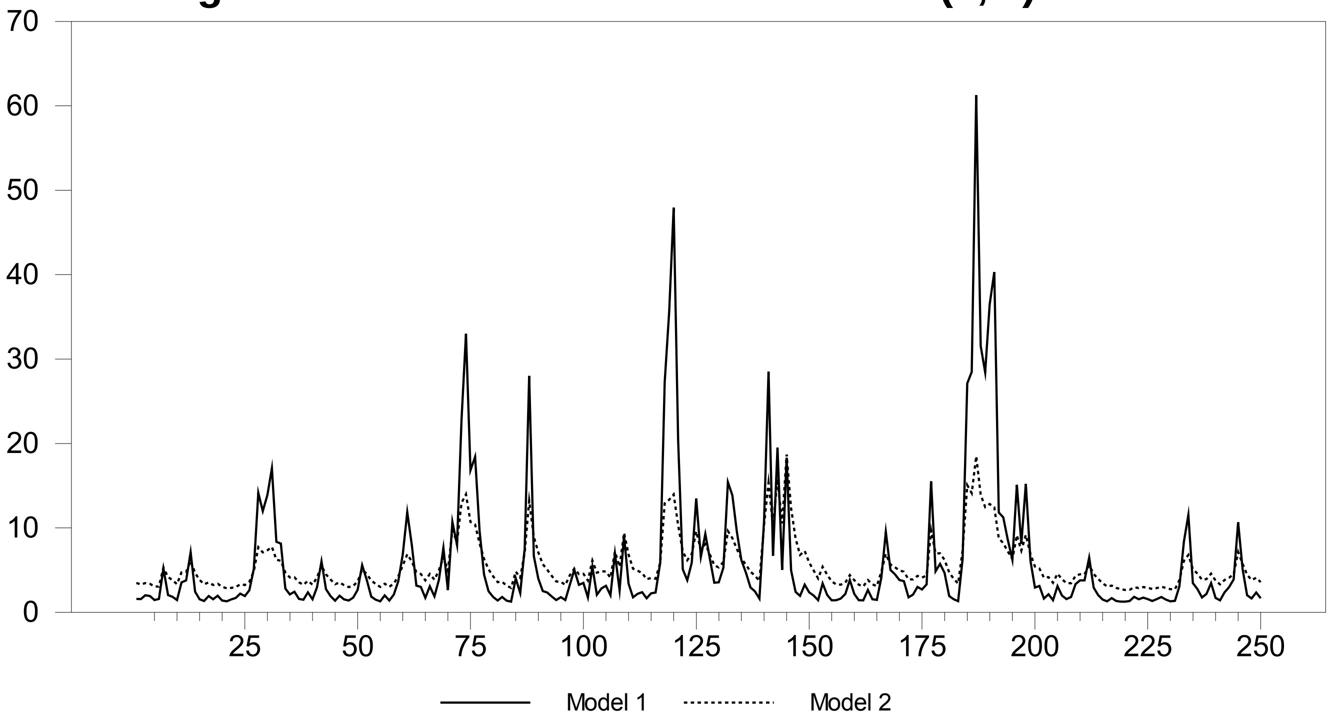
$$E_t h_{t+j} = \alpha_0 + (\alpha_1 + \beta_1)E_t h_{t+j-1}$$

$$E_t h_{t+j} = \alpha_0[1 + (\alpha_1 + \beta_1) + (\alpha_1 + \beta_1)^2 + \ldots + (\alpha_1 + \beta_1)^{j-1}] + (\alpha_1 + \beta_1)^j h_t$$

Volatility Persistence

Large values of both α_1 and β_1 act to increase the conditional volatility but they do so in different ways.

Figure 3.10: Persistence in the GARCH(1, 1) Model



$$h_t = 1 + 0.6\varepsilon_{t-1}^2 + 0.2h_{t-1} \qquad\qquad h_t = 1 + 0.2\varepsilon_{t-1}^2 + 0.6h_{t-1}$$

Assessing the Fit

Standardized residuals:

$$\text{RSS}' = \sum_{t=1}^{T} v_t^2 \qquad\qquad \text{RSS}' = \sum_{t=1}^{T} (\varepsilon_t^2 / h_t)$$

$$\text{AIC}' = -2 \ln L + 2n$$
$$\text{SBC}' = -2\ln L + n \ln(T)$$

where *L* likelihood function and *n* is the number of estimated parameters.

Diagnostic Checks for Model Adequacy

- If there is any serial correlation in the standardized residuals--the $\{s_t\}$ sequence--the model of the mean is not properly specified.
- To test for remaining GARCH effects, form the Ljung–Box Q-statistics of the squared standardized residuals.

Section 8

MAXIMUM LIKELIHOOD ESTIMATION

Maximum Likelihood and a Regression

Under the usual normality assumption, the log likelihood of observation *t* is:

$$(1/2)\ln(2\pi) - (1/2)\ln 2\sigma^2 - \frac{1}{2\sigma^2}(y_t - \beta x_t)^2$$

With T independent observations:

$$\log L = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T}(y_t - \beta x_t)^2$$

We want to select β and σ^2 so as to maximize L

$$\hat{\sigma}^2 = \sum \varepsilon_t^2 / T \qquad \hat{\beta} = \sum x_t y_t / \sum x_t^2$$

The Likelihood Function with ARCH errors

For observation *t* $$L_t = \left(\frac{1}{\sqrt{2\pi h_t}} \right) \exp\left(\frac{-\varepsilon_t^2}{2h_t} \right)$$

$$\ln L = -\frac{T}{2} \ln(2\pi) - 0.5 \sum_{t=1}^{T} \ln h_t - 0.5 \sum_{t=1}^{T} (\varepsilon_t^2 / h_t)$$

Now substitute for ε_t and h_t

$$\ln L = -\frac{T-1}{2} \ln(2\pi) - 0.5 \sum_{t=2}^{T} \ln(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2) - \frac{1}{2} \sum_{t=2}^{T} [(y_t - \beta x_t)^2 / (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)]$$

There are no analytic solutions to the first-order conditions for a maximum.

Section 9

OTHER MODELS OF CONDITIONAL VARIANCE

IGARCH

The IGARCH Model: Nelson (1990) argued that constraining α_1 + β_1 to equal unity can yield a very parsimonious representation of the distribution of an asset's return.

$$E_t h_{t+1} = \alpha_0 + h_t$$

and

$$E_t h_{t+j} = j\alpha_0 + h_t$$

$$h_t = \alpha_0 + (1-\beta_1)\varepsilon_{t-1}^2 + \beta_1 L h_t$$

$$h_t = \alpha_0 /(1-\beta_1) + (1-\beta_1)\sum_{i=0}^{\infty} \beta_1^i \varepsilon_{t-1-i}^2$$

RiskMetrics

RiskMetrics assumes that the continually compounded daily return of a portfolio follows a conditional normal distribution.

The assumption is that: $r_t|I_{t-1} \sim N(0, h_t)$

$$h_t = \alpha(\varepsilon_{t-1})^2 + (1 - \alpha)(h_{t-1}) \; ; \; \alpha > 0.9$$

Note: (Sometimes r_{t-1} is used). This is an IGARCH without an intercept.

Suppose that a loss occurs when the price falls. If the probability is 5%, RiskMetrics uses $1.65h_{t+1}$ to measure the risk of the portfolio. The Value at Risk (VaR) is:

$$\text{VaR} = \text{Amount of Position} \times 1.65(h_{t+1})^{1/2} \text{ and for k days is}$$

$$\text{VaR}(k) = \text{Amount of Position} \times 1.65(k\, h_{t+1})^{1/2}$$

Why Do We Care About ARCH Effects?

1. We care about the higher moments of the distribution.

2. The estimates of the coefficients of the mean are not correctly estimated if there are ARCH errors. Consider

$$\ln L = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T}(y_t - \beta x_t)^2$$

$$\ln L = -\frac{T}{2}\ln(2\pi) - 0.5\sum_{t=1}^{T}\ln h_t - 0.5\sum_{t=1}^{T}[(y - \beta x_t)^2 / h_t]$$

3. We want to place conditional confidence intervals around our forecasts
(see next page)

Example from Tsay

Suppose $r_t = 0.00066 – 0.0247r_{t-2} + \varepsilon_t$
$h_t = 0.00000389 + 0.0799(\varepsilon_{t-1})^2 + 0.9073(h_{t-1})$

Given current and past returns, suppose:
$E_T(r_{T+1}) = 0.00071$
and
$E_T(h_{T+1}) = 0.0003211$

The 5% quantile is $0.00071 – 1.6449*(0.0003211)^{1/2} = -0.02877$

The VaR for a portfolio size of $10,000,000 with probability 0.05 is

($10,000,000)(-0.02877) = $287,700

i.e., with 95% chance, the potential loss of the portfolio is $287,700 or less.

Models with Explanatory Variables

To model the effects of 9/11 on stock returns, create a dummy variable D_t equal to 0 before 9/11 and equal to 1 thereafter. Let

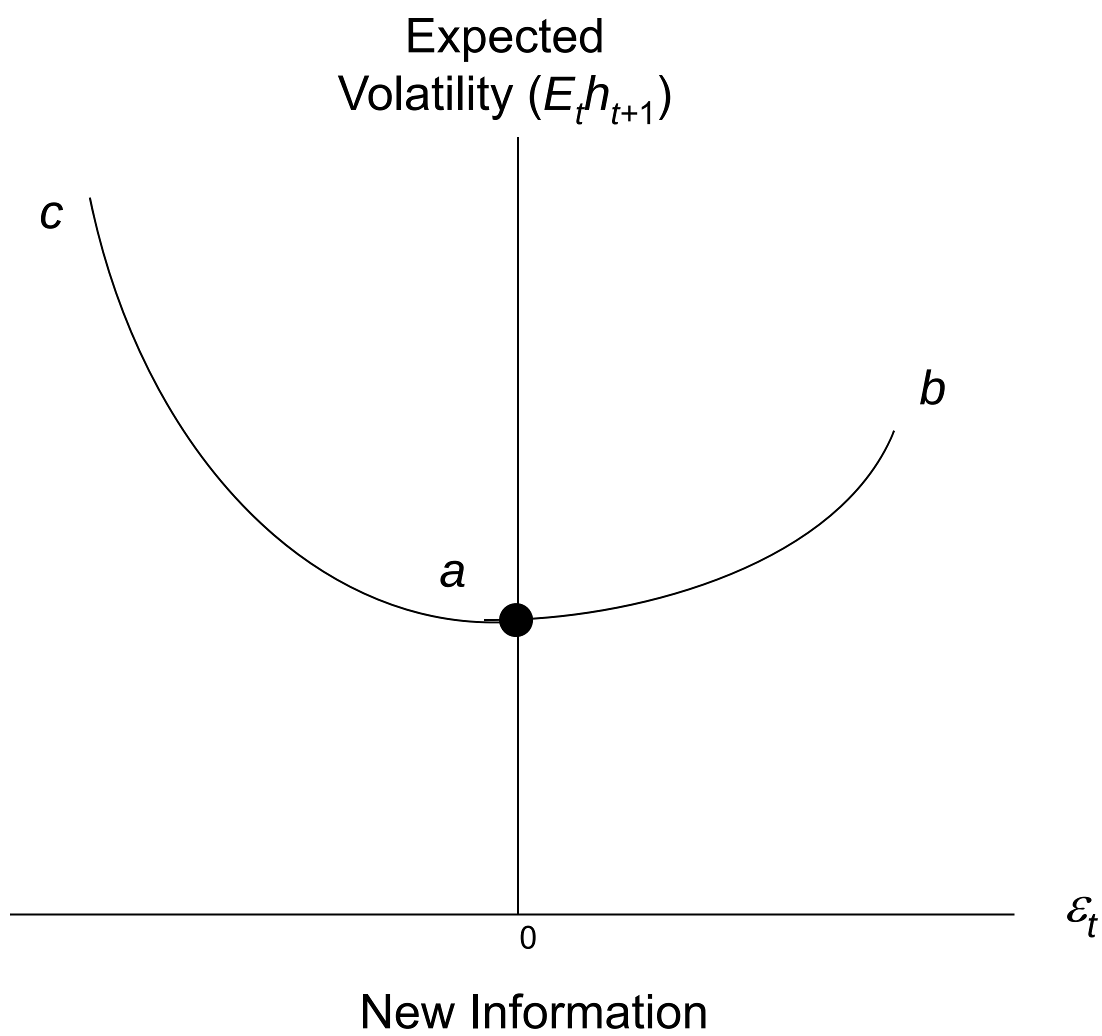
$$h_t = \alpha_0 + \alpha_1 + \beta_1 h_{t-1} + \gamma D_t$$

TARCH and EGARCH

- Glosten, Jaganathan and Runkle (1994) showed how to allow the effects of good and bad news to have different effects on volatility. Consider the threshold-GARCH (TARCH) process

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \lambda_1 d_{t-1} \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

Figure 3.11: The leverage effect



The EGARCH Model

$$\ln(h_t) = \alpha_0 + \alpha_1(\varepsilon_{t-1} / h_{t-1}^{0.5}) + \lambda_1 \,|\, \varepsilon_{t-1} / h_{t-1}^{0.5} \,| + \beta_1 \ln(h_{t-1})$$

1. The equation for the conditional variance is in log-linear form. Regardless of the magnitude of $\ln(h_t)$, the implied value of h_t can never be negative. Hence, it is permissible for the coefficients to be negative.

2. Instead of using the value of , the EGARCH model uses the level of standardized value of ε_{t-1} [i.e., ε_{t-1} divided by $(h_{t-1})^{0.5}$]. Nelson argues that this standardization allows for a more natural interpretation of the size and persistence of shocks. After all, the standardized value of ε_{t-1} is a unit-free measure.

3. The EGARCH model allows for leverage effects. If $\varepsilon_{t-1}/(h_{t-1})^{0.5}$ is positive, the effect of the shock on the log of the conditional variance is $\alpha_1 + \lambda_1$. If $\varepsilon_{t-1}/(h_{t-1})^{0.5}$ negative, the effect of the shock on the log of the conditional variance is $-\alpha_1 + \lambda_1$.

4. Although the EGARCH model has some advantages over the TARCH model, it is difficult to forecast the conditional variance of an EGARCH model.

Testing for Leverage Effects

1. If there are no leverage effects, the squared errors should be uncorrelated with the level of the error terms

$$s_t^2 = a_0 + a_1 s_{t-1} + a_2 s_{t-2} + \ldots$$

2. The Sign Bias test uses the regression equation of the form

$$s_t^2 = a_0 + a_1 d_{t-1} + \varepsilon_{1t}$$

where d_{t-1} is to 1 if $\varepsilon_{t-1} < 0$ and is equal to zero if $\varepsilon_{t-1} \geq 0$.

3. The more general test is

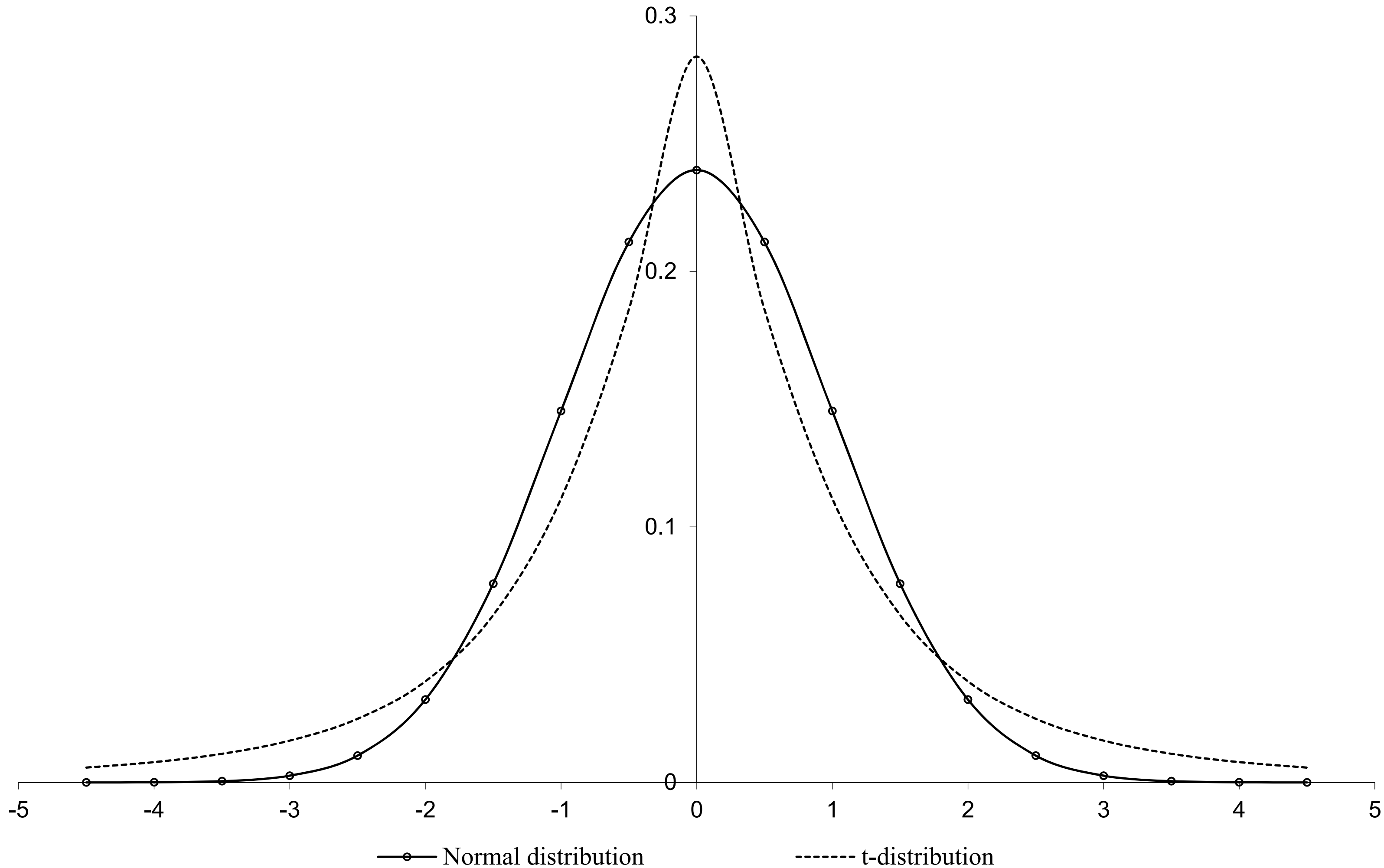
$$s_t^2 = a_0 + a_1 d_{t-1} + a_2 d_{t-1} s_{t-1} + a_3 (1 - d_{t-1}) s_{t-1} + \varepsilon_{1t}$$

$d_{t-1}s_{t-1}$ and $(1 - d_{t-1})s_{t-1}$ indicate whether the effects of positive and negative shocks also depend on their size. You can use an F-statistic to test the null hypothesis $a_1 = a_2 = a_3 = 0$.

Section 10

ESTIMATING THE NYSE U.S. 100 INDEX

Figure 3.12: Comparison of the Normal and *t* Distribuitions
(3 degrees of freedom)



The estimated model

$r_t = 0.043 + \varepsilon_t - 0.058 r_{t-1} - 0.038 r_{t-2}$ AIC = 9295.36, SBC = 9331.91

(2.82) (–3.00) (–1.91)

$h_t = 0.014 + 0.084\,(\varepsilon_t)^2 + 0.906\,h_{t-1}$

(4.91) (9.59) (98.31)

Instead, if we use a *t*-distribution, we obtain

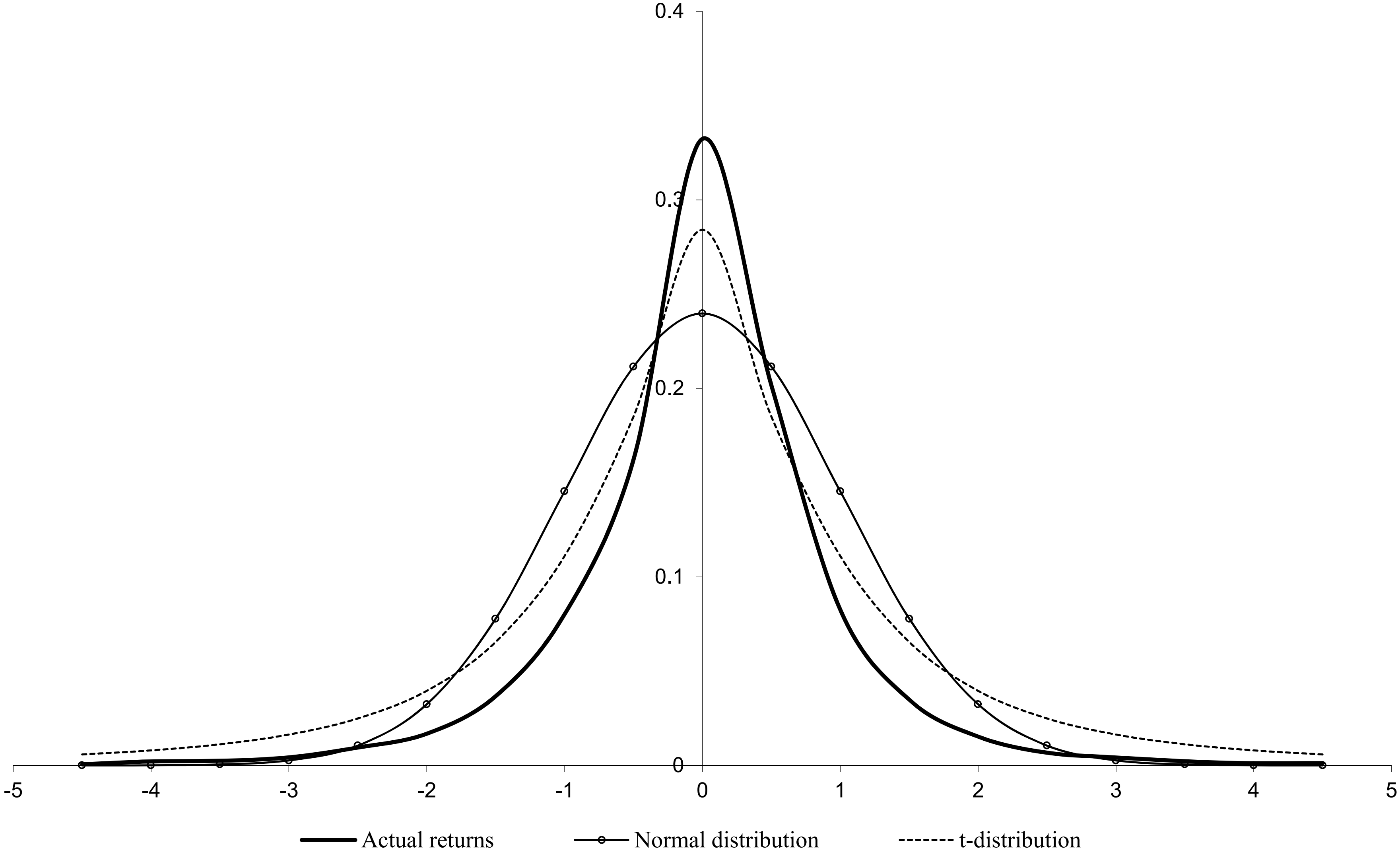
$r_t = 0.061 + \varepsilon_t - 0.062 r_{t-1} - 0.045 r_{t-2}$ AIC = 9162.72, SBC = 9205.37

(5.24) (–3.77) (–2.64)

$h_t = 0.009 + 0.089(\varepsilon_t)^2 + 0.909\,h_{t-1}$

(3.21) (8.58) (95.24)

Figure 3.13: Returns of the NYSE Index of 100 Stocks



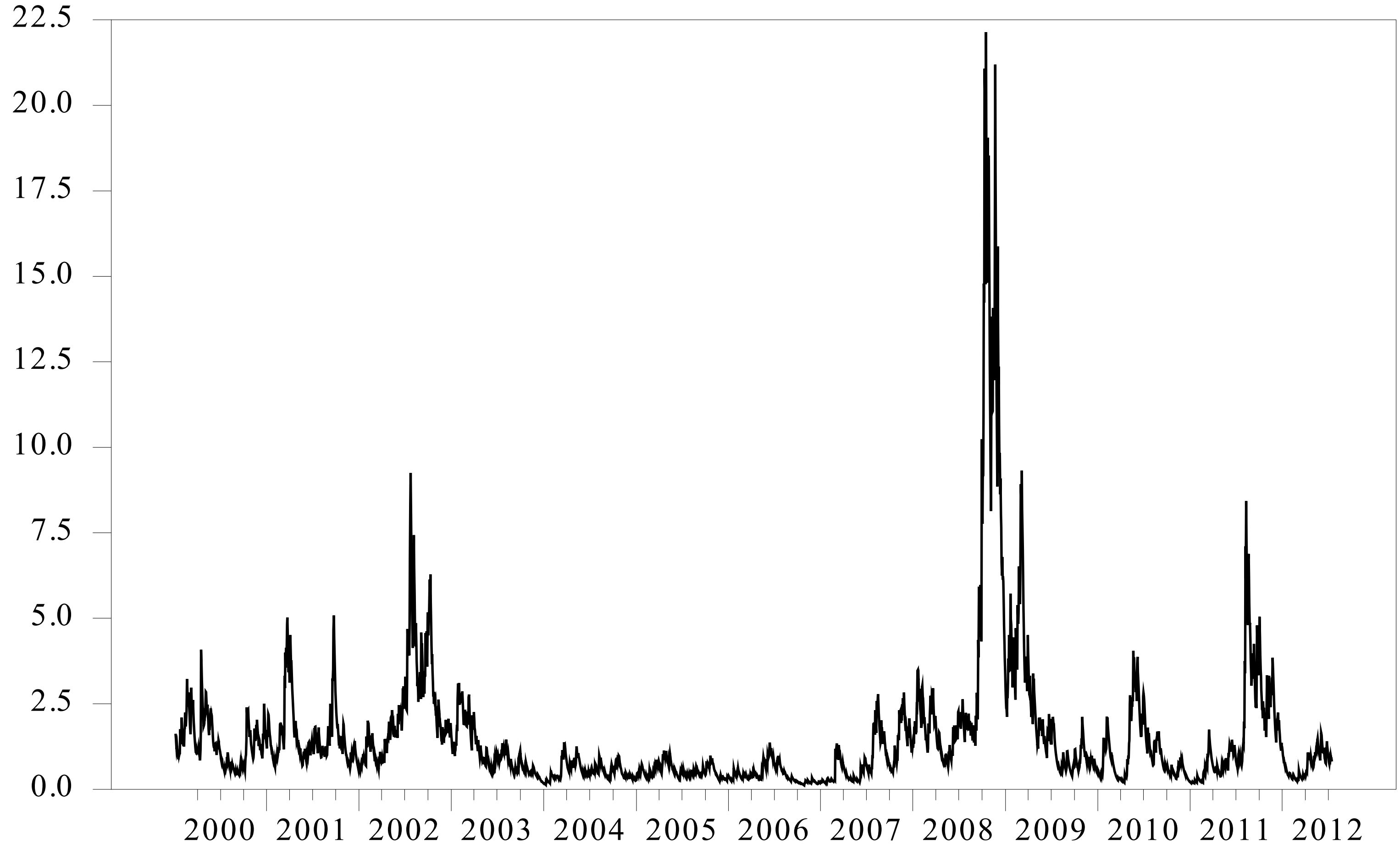


Figure 3.15: The Estimated Variance

Section 11

MULTIVARIATE GARCH

11. MULTIVARIATE GARCH

If you have a data set with several variables, it often makes sense to estimate the conditional volatilities of the variables simultaneously.

Multivariate GARCH models take advantage of the fact that the contemporaneous shocks to variables can be correlated with each other.

Equation-by-equation estimation is not efficient

Multivariate GARCH models allow for volatility spillovers in that volatility shocks to one variable might affect the volatility of other related variables

- Suppose there are just two variables, x_{1t} and x_{2t}. For now, we are not interested in the means of the series

- Consider the two error processes

$$\varepsilon_{1t} = v_{1t}(h_{11t})^{0.5}$$
$$\varepsilon_{2t} = v_{2t}(h_{22t})^{0.5}$$

- Assume $\text{var}(v_{1t}) = \text{var}(v_{2t}) = 1$, so that h_{11t} and h_{22t} are the conditional variances of ε_{1t} and ε_{2t}, respectively.

- We want to allow for the possibility that the shocks are correlated, denote h_{12t} as the conditional covariance between the two shocks. Specifically, let $h_{12t} = E_{t\text{-}1}\varepsilon_{1t}\varepsilon_{2t}$.

The VECH Model

A natural way to construct a multivariate GARCH(1, 1) is the *vech* model

$$h_{11t} = c_{10} + \alpha_{11} (\varepsilon_{1t-1})^2 + \alpha_{12}\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{13} (\varepsilon_{2t-1})^2 + \beta_{11}h_{11t-1} + \beta_{12}h_{12t-1} + \beta_{13}h_{22t-1}$$

$$h_{12t} = c_{20} + \alpha_{21} (\varepsilon_{1t-1})^2 + \alpha_{22}\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{23} (\varepsilon_{2t-1})^2 + \beta_{21}h_{11t-1} + \beta_{22}h_{12t-1} + \beta_{23}h_{22t-1}$$

$$h_{22t} = c_{30} + \alpha_{31} (\varepsilon_{1t-1})^2 + \alpha_{32}\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{33} (\varepsilon_{2t-1})^2 + \beta_{31}h_{11t-1} + \beta_{32}h_{12t-1} + \beta_{33}h_{22t-1}$$

The conditional variances (h_{11t} and h_{22t}) and covariance depend on their own past, the conditional covariance between the two variables (h_{12t}), the lagged squared errors, and the product of lagged errors ($\varepsilon_{1t-1}\varepsilon_{2t-1}$). Clearly, there is a rich interaction between the variables. After one period, a v_{1t} shock affects h_{11t}, h_{12t}, and h_{22t}.

ESTIMATION

Multivariate GARCH models can be very difficult to estimate. The number of parameters necessary can get quite large.

In the 2-variable case above, there are 21 parameters.

Once lagged values of $\{x_{1t}\}$ and $\{x_{2t}\}$ and/or explanatory variables are added to the mean equation, the estimation problem is complicated.

As in the univariate case, there is not an analytic solution to the maximization problem. As such, it is necessary to use numerical methods to find that parameter values that maximize the function L.

Since conditional variances are necessarily positive, the restrictions for the multivariate case are far more complicated than for the univariate case.

The results of the maximization problem must be such that every one of the conditional variances is always positive and that the implied correlation coefficients, $\rho_{ij} = h_{ij}/(h_{ii}h_{jj})^{0.5}$, are between –1 and +1.

The diagonal *vech*

- One set of restrictions that became popular in the early literature is the so-called diagonal *vech* model. The idea is to diagonalize the system such that h_{ijt} contains only lags of itself and the cross products of $\varepsilon_{it}\varepsilon_{jt}$. For example, the diagonalized version of (3.42) – (3.44) is

$$h_{11t} = c_{10} + \alpha_{11}(\varepsilon_{1t-1})^2 + \beta_{11}h_{11t-1}$$
$$h_{12t} = c_{20} + \alpha_{22}\varepsilon_{1t-1}\varepsilon_{2t-1} + \beta_{22}h_{12t-1}$$
$$h_{22t} = c_{30} + \alpha_{33}(\varepsilon_{2t-1})^2 + \beta_{33}h_{22t-1}$$

- Given the large number of restrictions, model is relatively easy to estimate.
- Each conditional variance is equivalent to that of a univariate GARCH process and the conditional covariance is quite parsimonious as well.
- The problem is that setting all $\alpha_{ij} = \beta_{ij} = 0$ (for $i \neq$ j) means that there are no interactions among the variances. A ε_{1t-1} shock, for example, affects h_{11t} and h_{12t}, but does not affect the conditional variance h_{2t}.

THE BEKK

•Engle and Kroner (1995) popularized what is now called the BEK (or BEKK) model that ensures that the conditional variances are positive. The idea is to force all of the parameters to enter the model via quadratic forms ensuring that all the variances are positive. Although there are several different variants of the model, consider the specification

$$H_t = C'C + A'\varepsilon_{t-1}\varepsilon_{t-1}'A + B'H_{t-1}B$$

where for the 2-variable case

$$C = \begin{bmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{bmatrix};\ A = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix};\ B = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix}$$

If you perform the indicated matrix multiplications you will find

$$h_{11t} = (c_{11}^2 + c_{12}^2) + (\alpha_{11}^2\varepsilon_{1t-1}^2 + 2\alpha_{11}\alpha_{21}\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{21}^2\varepsilon_{2t-1}^2) + (\beta_{11}^2 h_{11t-1} + 2\beta_{11}\beta_{21}h_{12t-1} + \beta_{21}^2 h_{22t-1})$$

THE BEK II

- In general, h_{ijt} will depend on the squared residuals, cross-products of the residuals, and the conditional variances and covariances of all variables in the system.
  - The model allows for shocks to the variance of one of the variables to "spill-over" to the others.
  - The problem is that the BEK formulation can be quite difficult to estimate. The model has a large number of parameters that are not globally identified. Changing the signs of all elements of *A*, *B* or *C* will have effects on the value of the likelihood function. As such, convergence can be quite difficult to achieve.
-

The BEKK (and Vech) as a VAR

$$h_{11t} = (c_{11}^2 + c_{12}^2) + (\alpha_{11}^2\varepsilon_{1t-1}^2 + 2\alpha_{11}\alpha_{21}\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{21}^2\varepsilon_{2t-1}^2) + (\beta_{11}^2 h_{11t-1} + 2\beta_{11}\beta_{21}h_{12t-1} + \beta_{21}^2 h_{22t-1})$$

$$h_{12t} = c_{12}(c_{11} + c_{22}) + \alpha_{12}\alpha_{11}\varepsilon_{1t-1}^2 + (\alpha_{11}\alpha_{22} + \alpha_{12}\alpha_{21})\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{21}\alpha_{22}\varepsilon_{2t-1}^2$$

$$+\beta_{11}\beta_{12}h_{11t-1} + (\beta_{11}\beta_{22} + \beta_{12}\beta_{21})h_{12t-1} + \beta_{21}\beta_{22}h_{22t-1}$$

$$h_{22t} = (c_{22}^2 + c_{12}^2) + (\alpha_{12}^2\varepsilon_{1t-1}^2 + 2\alpha_{12}\alpha_{22}\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{22}^2\varepsilon_{2t-1}^2)$$
$$+ (\beta_{12}^2 h_{11t-1} + 2\beta_{12}\beta_{22}h_{12t-1} + \beta_{22}^2 h_{22t-1})$$

Constant Conditional Correlations (CCC)

- As the name suggests, the (CCC)model restricts the correlation coefficients to be constant. As such, for each $i \neq j$, the CCC model assumes $h_{ijt} = \rho_{ij}(h_{iit}h_{jjt})^{0.5}$.

- In a sense, the CCC model is a compromise in that the variance terms need not be diagonalized, but the covariance terms are always proportional to $(h_{iit}h_{jjt})^{0.5}$. For example, a CCC model could consist of (3.42), (3.44) and

- $h_{12t} = \rho_{12}(h_{11t}h_{22t})^{0.5}$

- Hence, the covariance equation entails only one parameter instead of the 7 parameters appearing in (3.43).

EXAMPLE OF THE CCC MODEL

- Bollerslev (1990) examines the weekly values of the nominal exchange rates for five different countries--the German mark (DM), the French franc (FF), the Italian lira(IL), the Swiss franc (SF), and the British pound (BP)--relative to the U.S. dollar.
  - A five-equation system would be too unwieldy to estimate in an unrestricted form.
  - For the model of the mean, the log of each exchange rate series was modeled as a random walk plus a drift
  - $y_{it} = \mu_i + \varepsilon_{it}$ (3.45)
- *where* y_{it} is the percentage change in the nominal exchange rate for country *i*,
- Ljung-Box tests indicated each series of residuals did not contain any serial correlation

- Next, he tested the squared residuals for serial dependence. For example, for the British pound, the $Q(20)$-statistic has a value of 113.020; this is significant at any conventional level.

- Each series was estimated as a GARCH(1, 1) process.The specification has the form of (3.45) plus

$$h_{iit} = c_{i0} + \alpha_{ii}(\varepsilon_{\mathrm{i}t\text{-}1})^2 + \beta_{ii}h_{iit-1} \quad (i = 1, \ldots, 5)$$

$$h_{ijt} = \rho_{ij}(h_{iit}h_{jjt})^{0.5} \qquad (i \neq j)$$

- The model requires that only 30 parameters be estimated (five values of μ_i, the five equations for h_{iit} each have three parameters, and ten values of the ρ_{ij}).
- As in a seemingly unrelated regression framework, the system-wide estimation provided by the CCC model captures the contemporaneous correlation between the various error terms.

RESULTS

- The estimated correlations for the period during which the European Monetary System (EMS) prevailed are

	DM	FF	IL	SF
FF	0.932			
IL	0.886	0.876		
SW	0.917	0.866	0.816	
BP	0.674	0.678	0.622	0.635

It is interesting that correlations among continental European currencies were all far greater than those for the pound. Moreover, the correlations were much greater than those of the pre–EMS period. Clearly, EMS acted to keep the exchange rates of Germany, France, Italy and Switzerland tightly in line prior to the introduction of the Euro.

The file labeled EXRATES(DAILY).XLS contains the 2342 daily values of the Euro, British pound, and Swiss franc over the Jan. 3, 2000 – Dec. 23, 2008 period. Denote the U.S. dollar value of each of these nominal exchange rates as e_{it} where i = EU, BP and SW.

Construct the logarithmic change of each nominal exchange rate as $y_{it} = \log(e_{it}/e_{it-1})$. Although the residual autocorrelations are all very small in magnitude, a few are statistically significant. For example, the autocorrelations for the Euro are

ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6
0.036	–0.004	–0.004	0.063	0.001	–0.036

With $T = 2342$, the value of ρ_4 is statistically significant and the value of the Ljung-Box $Q(4)$ statistic is 12.37. Nevertheless, most researchers would not attempt to model this small value of the 4-*th* lag. Moreover, the SBC always selects models with no lagged changes in the mean equation.

For the second step, you should check the squared residuals for the presence of GARCH errors. Since we are using daily data (with a five-day week), it seems reasonable to begin using a model of the form The sample values of the *F*-statistics for the null hypothesis that $\alpha_1 = \ldots = \alpha_5 = 0$ are 43.36, 89.74, and 20.96 for the Euro, BP and SW, respectively. Since all of these values are highly significant, it is possible to conclude that all three series exhibit GARCH errors.

$$\hat{\varepsilon}_t^2 = \alpha_0 + \sum_{i=1}^{5} \alpha_i \hat{\varepsilon}_{t-5}^2$$

The sample values of the *F*-statistics for the null hypothesis that $\alpha_1 = \ldots = \alpha_5 = 0$ are 43.36, 89.74, and 20.96 for the Euro, BP and SW, respectively. Since all of these values are highly significant, it is possible to conclude that all three series exhibit GARCH errors.

If you estimate the three series as GARCH(1, 1) process using the CCC restriction, you should find the results reported in Table 3.1.

Table 3.1: The *CCC* Model of Exchange Rates

	c	α_1	β_1
Euro	$1.32\text{x}10^{-}$	0.047	0.951
	(2.44)	(10.79)	(240.91)
Pounc	$2.42\text{x}10^{-}$	0.040	0.953
	(3.28)	(7.71)	(149.15)
Franc	$2.16\text{x}10^{-}$	0.059	0.940
	(2.57)	(12/82)	(215.36)

If we let the numbers 1, 2, and 3 represent the euro, pound, and franc, the correlations are ρ_{12} = 0.68, ρ_{13} = 0.87, and ρ_{23} = 0.60. As in Bollerslev's paper, the pound and the franc continue to have the lowest correlation coefficient.

By way of contrast, it is instructive to estimate the model using the diagonal *vech* specification such that each variance and covariance is estimated separately. The estimation results are given in Table 3.2.

	h_{11t}	h_{12t}	h_{13t}	h_{22t}	h_{23t}	h_{33t}
c	$4.01\text{x}10^{-7}$	$2.50\ \text{x}10^{-7}$	$4.45\ \text{x}10^{-7}$	$2.62\ \text{x}10^{-7}$	$2.32\ \text{x}10^{-7}$	$5.88\ \text{x}10^{-7}$
	(18.47)	(6.39)	(33.82)	(4.31)	(6.39)	(10.79)
α_1	0.047	0.035	0.047	0.037	0.033	0.050
	(14.51)	(11.89)	(14.97)	(9.59)	(12.01)	(14.07)
β_1	0.946	0.956	0.945	0.956	0.959	0.941
	(319.44)	(268.97)	(339.91)	(205.04)	(309.29)	(270.55)

Now, the correlation coefficients are time varying. For example, the correlation coefficient between the pound and the franc is given by $h_{23t}/(h_{22t}h_{33t})^{0.5}$.

The time path of this correlation coefficient is shown as the solid line in Figure 3.16.

Although the correlation does seem to fluctuate around 0.64 (the value found by the CCC method), there are substantial departures from this average value.

Beginning in mid-2006, the correlation between the pound and the franc began a long and steady decline ending in March of 2008. The correlation increased with fears of a U.S. recession and then sharply fell with the onset on the U.S. financial crisis in the Fall of 2008.

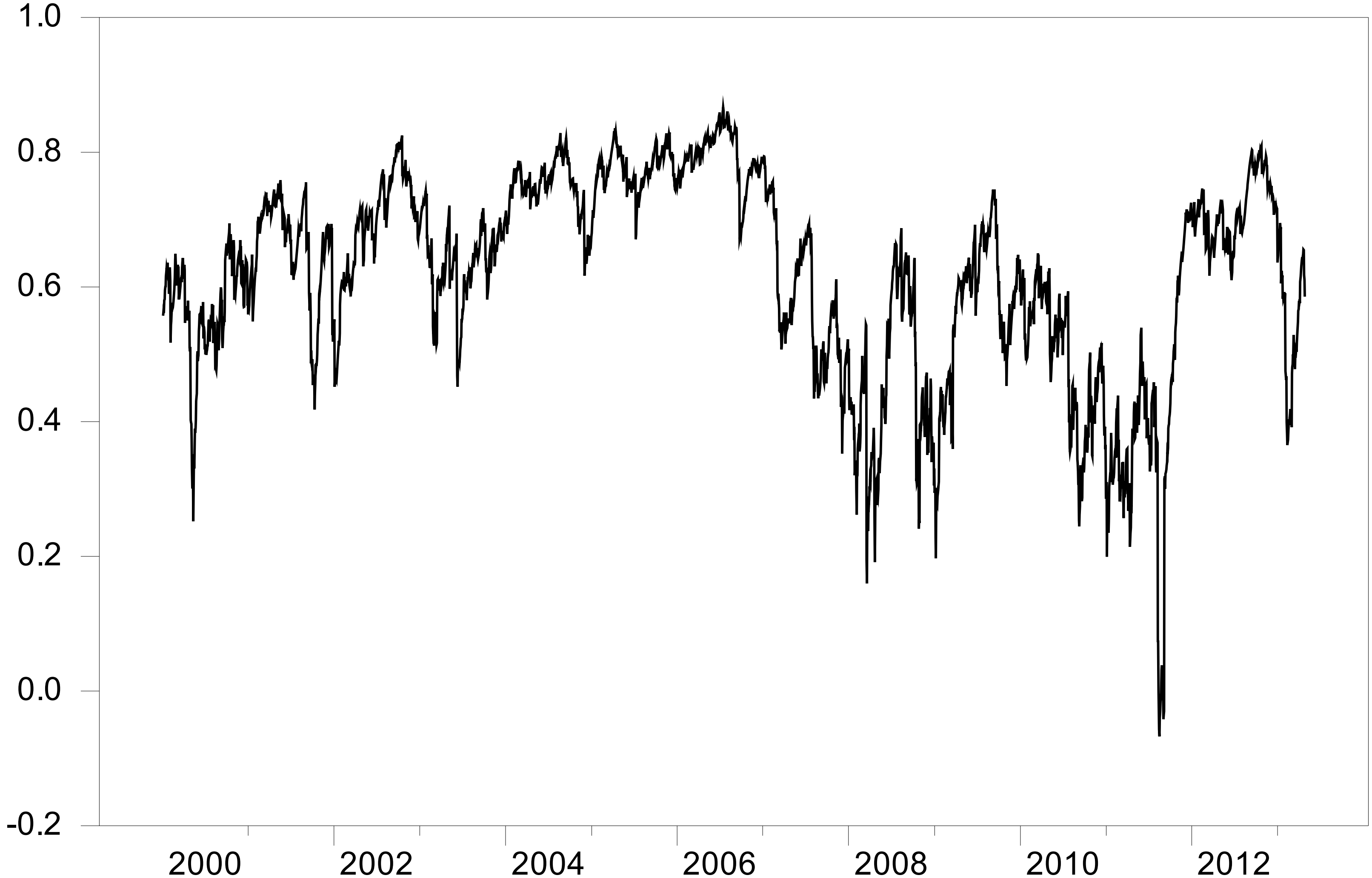


Figure 3.16: Pound/Franc Correlation from the Diagonal vech

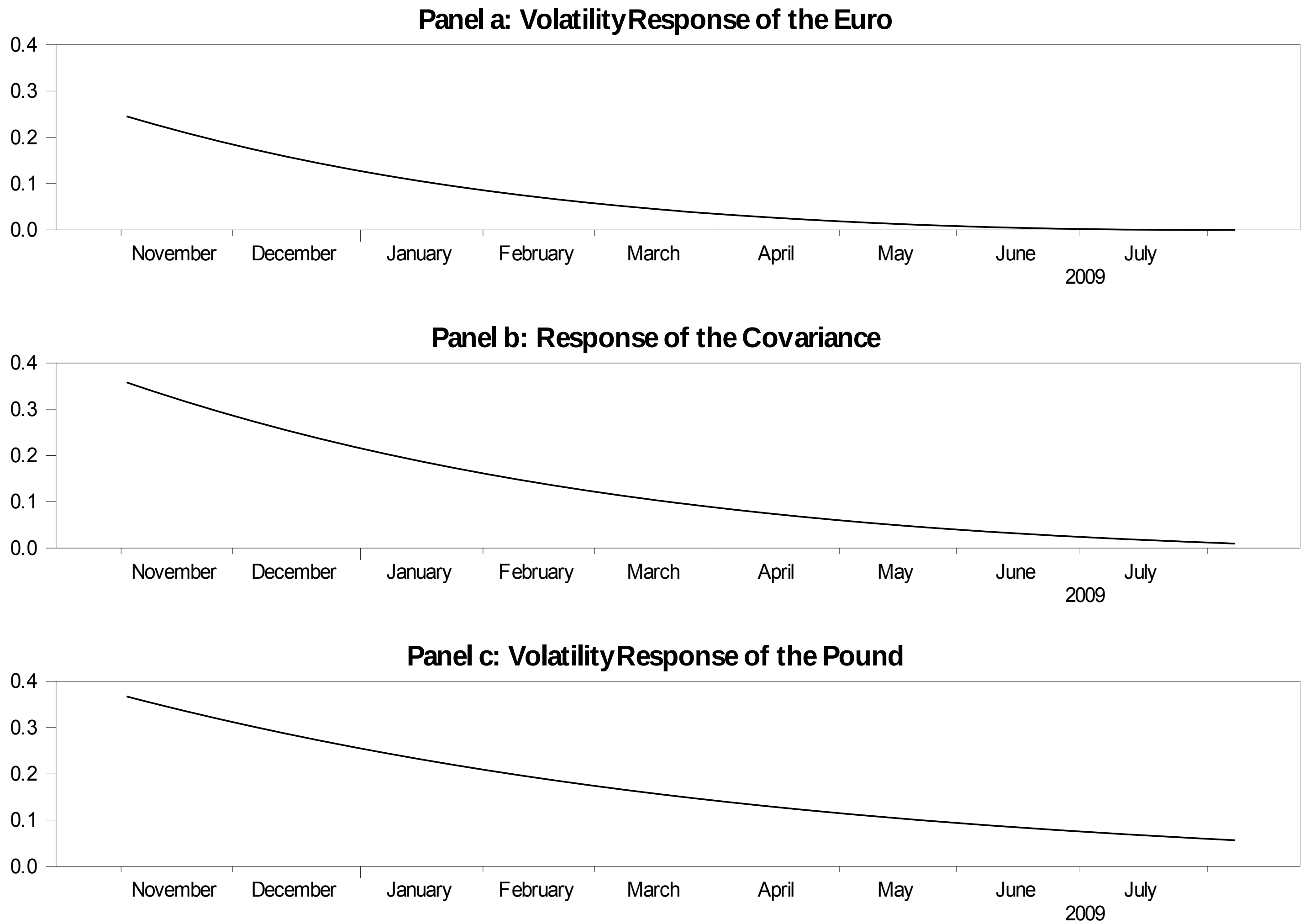


Figure 3.17 Variance Impulse Responses from Oct. 29, 2008

Appendix: The Log Likelihood Function

$$L_t = \frac{1}{2\pi\sqrt{h_{11}h_{22}(1-\rho_{12}^2)}}\exp\left[-\frac{1}{2(1-\rho_{12}^2)}\left(\frac{\varepsilon_{1t}^2}{h_{11}} + \frac{\varepsilon_{2t}^2}{h_{22}} - \frac{2\rho_{12}\varepsilon_{1t}\varepsilon_{2t}}{(h_{11}h_{22})^{0.5}}\right)\right]$$

where ρ_{12} is the correlation coefficient between ε_{1t} and ε_{2t};
$\rho_{12} = h_{12}/(h_{11}h_{22})^{0.5}$.

Now define

$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{bmatrix}$$

$$L_t = \frac{1}{2\pi|H|^{1/2}}\exp\left[-\frac{1}{2}\varepsilon_t' H^{-1}\varepsilon_t\right]$$

where ε_t = (ε_{1t}, ε_{2t})′, and | H | is the determinant of H.

Now, suppose that the realizations of $\{\varepsilon_t\}$ are independent, so that the likelihood of the joint realizations of ε_1, ε_2, … ε_T is the product in the individual likelihoods. Hence, if all have the same variance, the likelihood of the joint realizations is

$$L = \prod_{t=1}^{T} \frac{1}{2\pi |H|^{1/2}} \exp\left[-\frac{1}{2} \varepsilon_t' H^{-1} \varepsilon_t \right]$$

$$\ln L = -\frac{T}{2} \ln(2\pi) - \frac{T}{2} \ln |H| - \frac{1}{2} \sum_{t=1}^{T} \varepsilon_t' H^{-1} \varepsilon_t$$

MULTIVARIATE GARCH MODELS

For the 2-variable

$$L = \prod_{t=1}^{T} \frac{1}{2\pi |H_t|^{1/2}} \exp\left[-\frac{1}{2} \varepsilon_t' H_t^{-1} \varepsilon_t \right]$$

$$H_t = \begin{bmatrix} h_{11t} & h_{12t} \\ h_{12t} & h_{22t} \end{bmatrix}$$

$$\ln L = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^{T} (\ln |H_t| + \varepsilon_t' H_t^{-1} \varepsilon_t)$$

The form of the likelihood function is identical for models with *k* variables. In such circumstances, *H* is a symmetric *k* x *k* matrix, ε_t is a *k* x 1 column vector, and the constant term (2π) is raised to the power *k*.

The vech Operator

The *vech* operator transforms the upper (lower) triangle of a symmetric matrix into a column vector. Consider the symmetric covariance matrix

$$H_t = \begin{bmatrix} h_{11t} & h_{12t} \\ h_{12t} & h_{22t} \end{bmatrix} \qquad vech(H_t) = [\, h_{11t}, h_{12t}, h_{22t} \,]'$$

Now consider $\varepsilon_t = [\varepsilon_{1t}, \varepsilon_{2t}]'$. The product $\varepsilon_t \varepsilon_t' = [\varepsilon_{1t}, \varepsilon_{2t}]'[\varepsilon_{1t}, \varepsilon_{2t}]$ is

$$\begin{bmatrix} \varepsilon_{1t}^2 & \varepsilon_{1t}\varepsilon_{2t} \\ \varepsilon_{1t}\varepsilon_{2t} & \varepsilon_{1t}^2 \end{bmatrix}$$

$$vech(\varepsilon_t \varepsilon_t)' = \left[\varepsilon_{1t}^2, \varepsilon_{1t}\varepsilon_{2t}, \varepsilon_{2t}^2 \right]'$$

If we now let $C = [c_1, c_2, c_3]'$, A = the 3 x 3 matrix with elements α_{ij}, and B = the 3 x 3 matrix with elements β_{ij}, we can write

$$vech(H_t) = C + A\,vech(\varepsilon_{t-1}\varepsilon_{t-1}') + Bvech(H_{t-1})$$

it should be clear that this is precisely the system represented by (3.42) – (3.44). The diagonal *vech* uses only the diagonal elements of A and B and sets all values of $\alpha_{ij} = \beta_{ij} = 0$ for $i \neq j$.

Constant Conditional Correlations

$$H_t = \begin{bmatrix} h_{11t} & \rho_{12}(h_{11t}h_{22t})^{0.5} \\ \rho_{12}(h_{11t}h_{22t})^{0.5} & h_{22t} \end{bmatrix}$$

Now, if h_{11t} and h_{22t} are both GARCH(1, 1) processes, there are seven parameters to estimate (the six values of c_i, α_{ii} and β_{ii} and ρ_{12}).

Dynamic Conditional Correlations

- **STEP 1**: Use Bollerslev's CCC model to obtain the GARCH estimates of the variances and the standardized residuals
- **STEP 2:** Use the standardized residuals to estimate the conditional covariances.
  - Create the correlations by smoothing the series of standardized residuals obtained from the first step.
  - Engle examines several smoothing methods. The simplest is the exponential smoother $q_{ijt} = (1 - \lambda)s_{it}s_{jt} + \lambda q_{ijt\text{-}1}$ for $\lambda < 1$.
  - Hence, each $\{q_{iit}\}$ series is an exponentially weighted moving average of the cross-products of the standardized residuals.
  - The dynamic conditional correlations are created from the q_{ijt} as $\rho_{ijt} = q_{ijj}/(q_{ijt}q_{jjt})^{0.5}$