Advanced Topics in Time Series Econometrics Using R[1]

ZONGWU CAI[a,b]

E-mail address: **zcai@uncc.edu**

[a]Department of Mathematics & Statistics and Department of Economics,
University of North Carolina, Charlotte, NC 28223, U.S.A.

[b]Wang Yanan Institute for Studies in Economics, Xiamen University, China

July 18, 2007

---

[1]This manuscript may be printed and reproduced for individual or instructional use, but may not be printed for commercial purposes.

Contents

List of Tables

List of Figures

Chapter 1

Package R and Simple Applications

1.1 Computational Toolkits

When you work with large data sets, messy data handling, models, etc, you need to choose the computational tools that are useful for dealing with these kinds of problems. There are **"menu driven systems"** where you click some buttons and get some work done - but these are useless for anything nontrivial. To do serious economics and finance in the modern days, you have to write computer programs. And this is true of any field, for example, applied econometrics, empirical macroeconomics - and not just of "computational finance" which is a hot buzzword recently.

The **question** is how to choose the computational tools. According to Ajay Shah (December 2005), you should pay attention to three elements: **price, freedom, elegant** and **powerful computer science**, and **network effects**. Low price is better than high price. Price = 0 is obviously best of all. Freedom here is in many aspects. A good software system is one that does not tie you down in terms of hardware/OS, so that you are able to keep moving. Another aspect of freedom is in working with colleagues, collaborators and students. With commercial software, this becomes a problem, because your colleagues may not have the same software that you are using. Here free software really wins spectacularly. Good practice in research involves a great accent on reproducibility. Reproducibility is important both so as to avoid mistakes, and because the next person working in your field should be standing on your shoulders. This requires an ability to release code. This is only possible with free software. Systems like **SAS** and **Gauss** use archaic computer science. The code is inelegant. The language is not powerful. In this day and age, writing **C** or **Fortran** by hand is "too low level". Hell, with **Gauss**, even a minimal ting like online help is tawdry.

One prefers a system to be built by people who know their computer science - it should be an elegant, powerful language. All standard CS knowledge should be nicely in play to give you a gorgeous system. Good computer science gives you more productive humans. Lots of economists use **Gauss**, and give out **Gauss** source code, so there is a network effect in favor of **Gauss**. A similar thing is right now happening with statisticians and **R**.

Here I cite comparisons among most commonly used packages (see Ajay Shah (December 2005)); see the web site at
http://www.mayin.org/ajayshah/COMPUTING/mytools.html.

R is a very convenient programming language for doing statistical analysis and Monte Carol simulations as well as various applications in quantitative economics and finance. Indeed, we prefer to think of it of an environment within which statistical techniques are implemented. I will teach it at the introductory level, but NOTICE that you will have to learn **R** on your own. Note that about 97% of commands in **S-PLUS** and **R** are same. In particular, for analyzing time series data, **R** has a lot of bundles and packages, which can be downloaded for free, for example, at **http://www.r-project.org/**.

R, like **S**, is designed around a true computer language, and it allows users to add additional functionality by defining new functions. Much of the system is itself written in the **R** dialect of **S**, which makes it easy for users to follow the algorithmic choices made. For computationally-intensive tasks, **C**, **C++** and **Fortran** code can be linked and called at run time. Advanced users can write **C** code to manipulate **R** objects directly.

1.2 How to Install R ?

(1) go to the web site **http://www.r-project.org/**;
(2) click **CRAN**;
(3) choose a site for downloading, say **http://cran.cnr.Berkeley.edu**;
(4) click **Windows (95 and later)**;
(5) click **base**;
(6) click **R-2.5.1-win32.exe** (Version of 06-28-2007) to save this file first and then run it to install.
The basic R is installed into your computer. If you need to install other packages, you need

to do the followings:

(7) After it is installed, there is an icon on the screen. Click the icon to get into **R**;

(8) Go to the top and find **packages** and then click it;

(9) Go down to **Install package(s)...** and click it;

(10) There is a new window. Choose a location to download the packages, say **USA(CA1)**, move mouse to there and click OK;

(11) There is a new window listing all packages. You can select any one of packages and click OK, or you can select all of them and then click OK.

1.3 Data Analysis and Graphics Using R – An Introduction (109 pages)

See the file **r-notes.pdf** (109 pages) which can be downloaded from
http://www.math.uncc.edu/˜ zcai/r-notes.pdf.
I encourage you to download this file and learn it by yourself.

1.4 CRAN Task View: Empirical Finance

This CRAN Task View contains a list of packages useful for empirical work in Finance, grouped by topic. Besides these packages, a very wide variety of functions suitable for empirical work in Finance is provided by both the basic **R** system (and its set of recommended core packages), and a number of other packages on the Comprehensive **R** Archive Network (CRAN). Consequently, several of the other CRAN Task Views may contain suitable packages, in particular the **Econometrics** Task View. The web site is
http://cran.r-project.org/src/contrib/Views/Finance.html

1. **Standard regression models:** Linear models such as ordinary least squares (**OLS**) can be estimated by **lm**() (from by the stats package contained in the basic R distribution). Maximum Likelihood (ML) estimation can be undertaken with the **optim**() function. Non-linear least squares can be estimated with the **nls**() function, as well as with **nlme**() from the **nlme** package. For the linear model, a variety of regression diagnostic tests are provided by the **car**, **lmtest**, **strucchange**, **urca**, **uroot**, and sandwich packages. The **Rcmdr** and **Zelig** packages provide user interfaces that may be of interest as well.

2. **Time series:** Classical time series functionality is provided by the **arima**() and **KalmanLike**() commands in the basic R distribution. The **dse** packages provides a variety of more advanced estimation methods; **fracdiff** can estimate fractionally integrated series; **longmemo** covers related material. For **volatily** modeling, the standard **GARCH**(1,1) model can be estimated with the **garch**() function in the **tseries** package. Unit root and cointegration tests are provided by **tseries**, **urca** and **uroot**. The **Rmetrics** packages **fSeries** and **fMultivar** contain a number of estimation functions for ARMA, GARCH, long memory models, unit roots and more. The **ArDec** implements autoregressive time series decomposition in a Bayesian framework. The **dyn** and **dynlm** are suitable for dynamic (linear) regression models. Several packages provide wavelet analysis functionality: **rwt**, **wavelets**, **waveslim**, **wavethresh**. Some methods from chaos theory are provided by the package **tseriesChaos**.

3. **Finance:** The **Rmetrics** bundle comprised of the **fBasics**, **fCalendar**, **fSeries**, **fMultivar**, **fPortfolio**, **fOptions** and **fExtremes** packages contains a very large number of relevant functions for different aspect of empirical and computational finance. The **RQuantLib** package provides several option-pricing functions as well as some fixed-income functionality from the **QuantLib** project to **R**. The portfolio package contains classes for equity portfolio management.

4. **Risk Management:** The **VaR** package estimates Value-at-Risk, and several packages provide functionality for Extreme Value Theory models: **evd**, **evdbayes**, **evir**, **extRremes**, **ismec**, **POT**. The **mvtnorm** package provides code for multivariate Normal and t-distributions. The **Rmetrics** packages **fPortfolio** and **fExtremes** also contain a number of relevant functions. The **copula** and **fgac** packages cover multivariate dependency structures using copula methods.

5. **Data and Date Management:** The **its**, **zoo** and **fCalendar** (part of **Rmetrics**) packages provide support for irregularly-spaced time series. **fCalendar** also addresses calendar issues such as recurring holidays for a large number of financial centers, and provides code for high-frequency data sets.

   Related links:

* CRAN Task View: Econometrics. The web site is
  http://cran.cnr.berkeley.edu/src/contrib/Views/Econometrics.html
  or see the next section.

* Rmetrics by Diethelm Wuertz contains a wealth of **R** code for Finance. The web site is
  http://www.itp.phys.ethz.ch/econophysics/R/

* Quantlib is a C++ library for quantitative finance. The web site is
  http://quantlib.org/

* Mailing list: R Special Interest Group Finance

1.5 CRAN Task View: Computational Econometrics

Base **R** ships with a lot of functionality useful for computational econometrics, in particular in the stats package. This functionality is complemented by many packages on CRAN, a brief overview is given below. There is also a considerable overlap between the tools for econometrics in this view and finance in the **Finance** view. Furthermore, the **finance SIG** is a suitable mailing list for obtaining help and discussing questions about both computational finance and econometrics. The packages in this view can be roughly structured into the following topics. The web site is
http://cran.r-project.org/src/contrib/Views/Econometrics.html

1. **Linear regression models:** Linear models can be fitted (via OLS) with **lm**() (from stats) and standard tests for model comparisons are available in various methods such as **summary**() and **anova**(). Analogous functions that also support asymptotic tests (z instead of t tests, and Chi-squared instead of F tests) and plug-in of other covariance matrices are **coeftest**() and **waldtest**() in **lmtest**. Tests of more general linear hypotheses are implemented in **linear.hypothesis**() in car. HC and HAC covariance matrices that can be plugged into these functions are available in sandwich. The packages **car** and lmtest also provide a large collection of further methods for diagnostic checking in linear regression models.

2. **Microeconometrics:** Many standard micro-econometric models belong to the family of generalized linear models (GLM) and can be fitted by **glm**() from package stats. This

includes in particular logit and probit models for modelling choice data and poisson models for count data. Negative binomial GLMs are available via **glm.nb**() in package MASS from the **VR** bundle. Zero-inflated count models are provided in **zicounts**. Further over-dispersed and inflated models, including hurdle models, are available in package **pscl**. Bivariate poisson regression models are implemented in **bivpois**. Basic censored regression models (e.g., tobit models) can be fitted by **survreg**() in **survival**. Further more refined tools for microecnometrics are provided in **micEcon**. The package **bayesm** implements a Bayesian approach to microeconometrics and marketing. Inference for relative distributions is contained in package **reldist**.

3. **Further regression models:** Various extensions of the linear regression model and other model fitting techniques are available in base R and several CRAN packages. Nonlinear least squares modelling is available in **nls**() in package stats. Relevant packages include quantreg (quantile regression), **sem** (linear structural equation models, including two-stage least squares), **systemfit** (simultaneous equation estimation), **betareg** (beta regression), **nlme** (nonlinear mixed-effect models), **VR** (multinomial logit models in package **nnet**) and **MNP** (Bayesian multinomial probit models). The packages **Design** and **Hmisc** provide several tools for extended handling of (generalized) linear regression models.

4. **Basic time series infrastructure:** The class **ts** in package stats is R's standard class for regularly spaced time series which can be coerced back and forth without loss of information to **zooreg** from package **zoo**. **zoo** provides infrastructure for both regularly and irregularly spaced time series (the latter via the class "zoo") where the time information can be of arbitrary class. Several other implementations of irregular time series building on the "POSIXt" time-date classes are available in **its**, **tseries** and **fCalendar** which are all aimed particularly at finance applications (see the **Finance** view).

5. **Time series modelling:** Classical time series modelling tools are contained in the stats package and include **arima**() for ARIMA modelling and Box-Jenkins-type analysis. Furthermore stats provides **StructTS**() for fitting structural time series and **decompose**() and **HoltWinters**() for time series filtering and decomposition. For estimating VAR models, several methods are available: simple models can be fitted

by **ar**() in stats, more elaborate models are provided by **estVARXls**() in **dse** and a Bayesian approach is available in MSBVAR. A convenient interface for fitting dynamic regression models via OLS is available in **dynlm**; a different approach that also works with other regression functions is implemented in **dyn**. More advanced dynamic system equations can be fitted using **dse**. Unit root and cointegration techniques are available in **urca**, **uroot** and **tseries**. Time series factor analysis is available in **tsfa**.

6. **Matrix manipulations:** As a vector- and matrix-based language, base **R** ships with many powerful tools for doing matrix manipulations, which are complemented by the packages **Matrix** and **SparseM**.

7. **Inequality:** For measuring inequality, concentration and poverty the package **ineq** provides some basic tools such as Lorenz curves, Pen's parade, the Gini coefficient and many more.

8. **Structural change:** **R** is particularly strong when dealing with structural changes and **changepoints** in parametric models, see **strucchange** and **segmented**.

9. **Data sets:** Many of the packages in this view contain collections of data sets from the econometric literature and the package **Ecdat** contains a complete collection of data sets from various standard econometric textbooks. **micEcdat** provides several data sets from the Journal of Applied Econometrics and the Journal of Business & Economic Statistics data archives. Package **CDNmoney** provides Canadian monetary aggregates and **pwt** provides the Penn world table.

Related links:

* CRAN Task View: Finance. The web site is **http://cran.cnr.berkeley.edu/src/contrib/ Views/Finance.html** or see the above section.
* Mailing list: **R** Special Interest Group Finance
* A Brief Guide to **R** for Beginners in Econometrics. The web site is **http://people.su.se/~ ma/R_intro/**.

* **R** for Economists. The web site is
  http://www.mayin.org/ajayshah/KB/R/R_for_economists.html.

Chapter 2

Regression Models With Correlated Errors

2.1 Methodology

In many applications, the relationship between two time series is of major interest. The market model in finance is an example that relates the return of an individual stock to the return of a market index. The term structure of interest rates is another example in which the time evolution of the relationship between interest rates with different maturities is investigated. These examples lead to the consideration of a linear regression in the form $y_t = \beta_1 + \beta_2 x_t + e_t$, where y_t and x_t are two time series and e_t denotes the error term. The least squares (LS) method is often used to estimate the above model. If $\{e_t\}$ is a white noise series, then the LS method produces consistent estimates. In practice, however, it is common to see that the error term e_t is serially correlated. In this case, we have a regression model with time series errors, and the LS estimates of β_1 and β_2 may not be consistent and efficient.

Regression model with time series errors is widely applicable in economics and finance, but it is one of the most commonly misused econometric models because the serial dependence in e_t is often overlooked. It pays to study the model carefully. The standard method for dealing with correlated errors e_t in the regression model

$$y_t = \boldsymbol{\beta}^T \mathbf{z}_t + e_t \tag{2.1}$$

is to try to transform the errors e_t into uncorrelated ones and then apply the standard least squares approach to the transformed observations. For example, let $\mathbf{P}$ be an $n \times n$ matrix that transforms the vector $\mathbf{e} = (e_1, \cdots, e_n)^T$ into a set of independent identically distributed variables with variance σ^2. Then, the matrix version of (2.1) is

$$\mathbf{P}\,\mathbf{y} = \mathbf{P}\mathbf{Z}\,\boldsymbol{\beta} + \mathbf{P}\,\mathbf{e}$$

and proceed as before. Of course, the major problem is deciding on what to choose for $\mathbf{P}$ but in the time series case, happily, there is a reasonable solution, based again on time series ARMA models. Suppose that we can find, for example, a reasonable ARMA model for the residuals, say, for example, the ARMA$(p, 0, 0)$ model

$$e_t = \sum_{k=1}^{p} \phi_k\, e_t + w_t,$$

which defines a linear transformation of the correlated e_t to a sequence of uncorrelated w_t. We can ignore the problems near the beginning of the series by starting at $t = p$. In the ARMA notation, using the back-shift operator B, we may write

$$\phi(L)\, e_t = w_t, \tag{2.2}$$

where

$$\phi(L) = 1 - \sum_{k=1}^{p} \phi_k\, L^k \tag{2.3}$$

and applying the operator to both sides of (2.1) leads to the model

$$\phi(L)\, y_t = \phi(L)\, z_t + w_t, \tag{2.4}$$

where the $\{w_t\}$'s now satisfy the independence assumption. Doing ordinary least squares on the transformed model is the same as doing weighted least squares on the untransformed model. The only problem is that we do not know the values of the coefficients ϕ_k $(1 \le k \le p)$ in the transformation (2.3). However, if we knew the residuals e_t, it would be easy to estimate the coefficients, since (2.3) can be written in the form

$$e_t = \boldsymbol{\phi}^T\, \mathbf{e}_{t-1} + w_t, \tag{2.5}$$

which is exactly the usual regression model (2.1) with $\boldsymbol{\phi} = (\phi_1, \cdots, \phi_p)^T$ replacing $\boldsymbol{\beta}$ and $\mathbf{e}_{t-1} = (e_{t-1}, e_{t-2}, \cdots, e_{t-p})^T$ replacing $\mathbf{z}_t$. The above comments suggest a general approach known as the Cochran-Orcutt (1949) procedure for dealing with the problem of correlated errors in the time series context.

1. Begin by fitting the original regression model (2.1) by least squares, obtaining $\boldsymbol{\beta}$ and the residuals $e_t = y_t - \widehat{\boldsymbol{\beta}}^T \mathbf{z}_t$.

2. Fit an ARMA to the estimated residuals, say $\phi(L)\, e_t = \theta(L)\, w_t$.

3. Apply the ARMA transformation found to both sides of the regression equation (2.1) to obtain
$$\frac{\phi(L)}{\theta(L)}\, y_t = \boldsymbol{\beta}^T \frac{\phi(L)}{\theta(L)}\, \mathbf{z}_t + w_t.$$

4. Run an ordinary least squares on the transformed values to obtain the new $\boldsymbol{\beta}$.

5. Return to 2. if desired.

Often, one iteration is enough to develop the estimators under a reasonable correlation structure. In general, the Cochran-Orcutt procedure converges to the maximum likelihood or weighted least squares estimators.

Note that there is a function in **R** to compute the Cochrane-Orcutt estimator

```
arima(x, order = c(0, 0, 0),
        seasonal = list(order = c(0, 0, 0), period = NA),
        xreg = NULL, include.mean = TRUE, transform.pars = TRUE,
        fixed = NULL, init = NULL, method = c("CSS-ML", "ML", "CSS"),
        n.cond, optim.control = list(), kappa = 1e6)
```

by specifying “xreg=...”, where xreg is a vector or matrix of external regressors, which must have the same number of rows as “x”.

Example 3.1: The data shown in Figure 2.1 represent quarterly earnings per share for the American Company Johnson & Johnson from the from the fourth quarter of 1970 to the first quarter of 1980. We might consider an alternative approach to treating the Johnson and Johnson earnings series, assuming that $y_t = \log(x_t) = \beta_1 + \beta_2\, t + e_t$. In order to analyze the data with this approach, first we fit the model above, obtaining $\widehat{\beta}_1 = -0.6678(0.0349)$ and $\widehat{\beta}_2 = 0.0417(0.0071)$. The computed residuals $e_t = y_t - \widehat{\beta}_1 - \widehat{\beta}_2\, t$ can be computed easily, the ACF and PACF are shown in the top two panels of Figure 2.2. Note that the ACF and PACF suggest that a seasonal AR series will fit well and we show the ACF and PACF of these residuals in the bottom panels of Figure 2.2. The seasonal AR model is of the form

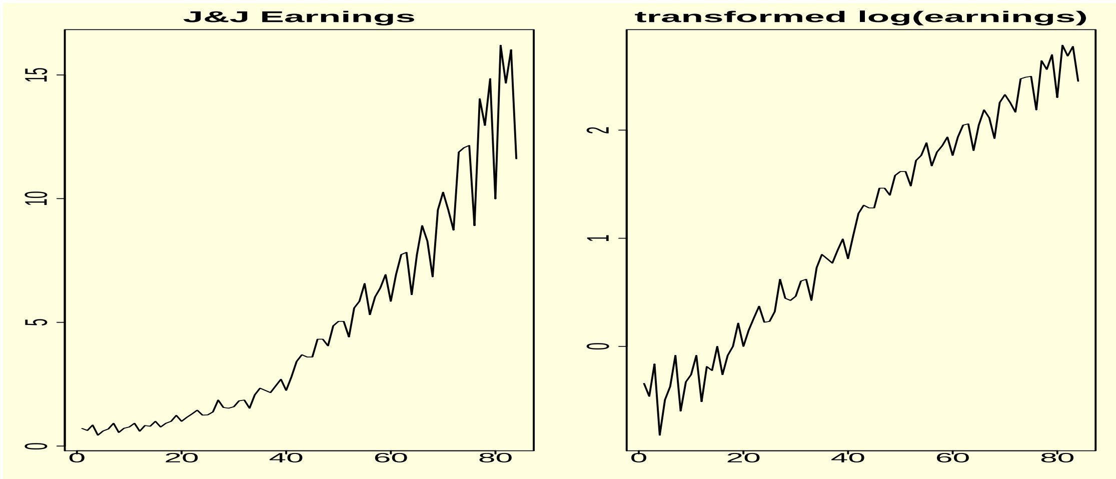

Figure 2.1: Quarterly earnings for Johnson & Johnson (4th quarter, 1970 to 1st quarter, 1980, left panel) with log transformed earnings (right panel).

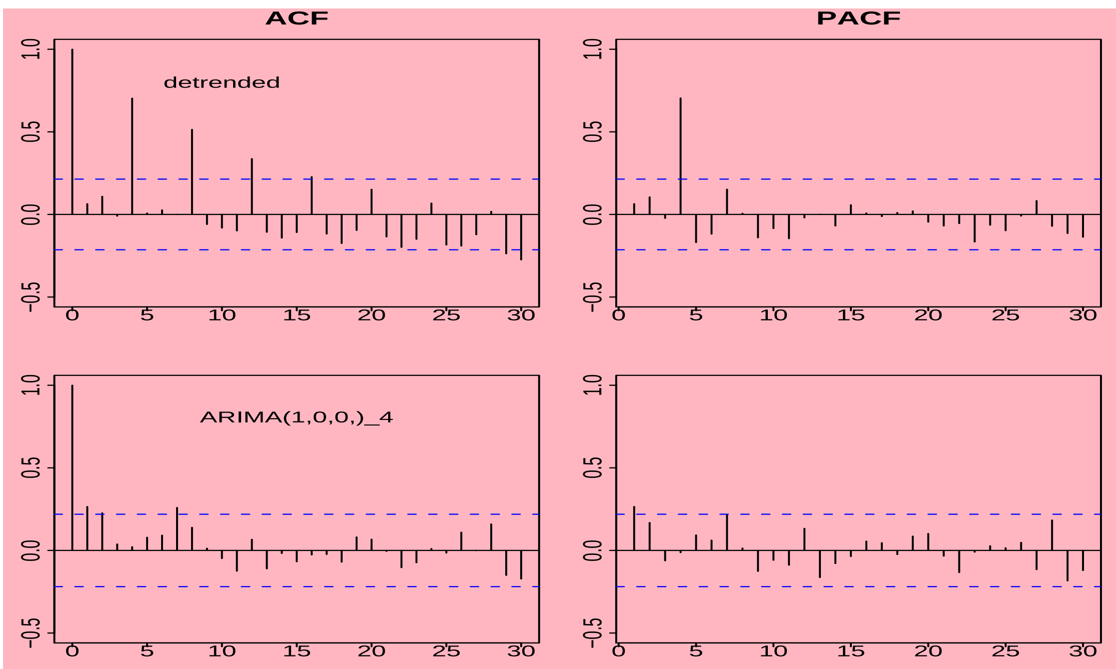

Figure 2.2: Autocorrelation functions (ACF) and partial autocorrelation functions (PACF) for the detrended log J&J earnings series (top two panels)and the fitted ARIMA$(0,0,0)\times(1,0,0)_4$ residuals.

$e_t = \Phi_1\, e_{t-4} + w_t$ and we obtain $\widehat{\Phi}_1 = 0.7614(0.0639)$, with $\widehat{\sigma}_w^2 = 0.00779$. Using these values, we transform y_t to

$$y_t - \widehat{\Phi}_1\, y_{t-4} = \beta_1(1-\widehat{\Phi}_1) + \beta_2[t - \widehat{\Phi}_1(t-4)] + w_t$$

using the estimated value $\widehat{\Phi}_1 = 0.7614$. With this transformed regression, we obtain the new estimators $\widehat{\beta}_1 = -0.7488(0.1105)$ and $\widehat{\beta}_2 = 0.0424(0.0018)$. The new estimator has the advantage of being unbiased and having a smaller generalized variance.

To forecast, we consider the original model, with the newly estimated $\widehat{\beta}_1$ and $\widehat{\beta}_2$. We obtain the approximate forecast for $y_{t+h}^t = \widehat{\beta}_1 + \widehat{\beta}_2(t+h) + e_{t+h}^t$ for the log transformed series, along with upper and lower limits depending on the estimated variance that only incorporates the prediction variance of e_{t+h}^t, considering the trend and seasonal autoregressive parameters as fixed. The narrower upper and lower limits (The figure is not presented here) are mainly a refection of a slightly better fit to the residuals and the ability of the trend model to take care of the nonstationarity.

Example 3:2: We consider the relationship between two U.S. weekly interest rate series: x_t: the 1-year Treasury constant maturity rate and y_t: the 3-year Treasury constant maturity rate. Both series have 1967 observations from January 5, 1962 to September 10, 1999 and are measured in percentages. The series are obtained from the Federal Reserve Bank of St Louis.

Figure 2.3 shows the time plots of the two interest rates with solid line denoting the 1-year rate and dashed line for the 3-year rate. The left panel of Figure 2.4 plots y_t versus

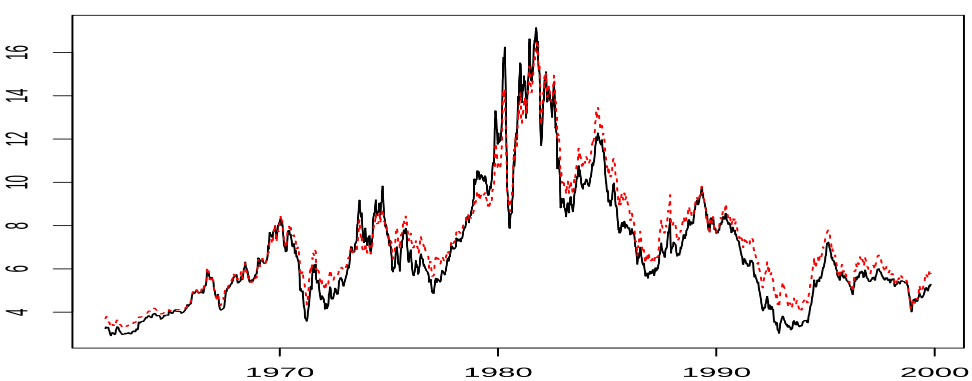

Figure 2.3: Time plots of U.S. weekly interest rates (in percentages) from January 5, 1962 to September 10, 1999. The solid line (black) is the Treasury 1-year constant maturity rate and the dashed line the Treasury 3-year constant maturity rate (red).

x_t, indicating that, as expected, the two interest rates are highly correlated. A naive way to describe the relationship between the two interest rates is to use the simple model, Model I: $y_t = \beta_1 + \beta_2\, x_t + e_t$. This results in a fitted model $y_t = 0.911 + 0.924\, x_t + e_t$, with $\widehat{\sigma}_e^2 = 0.538$

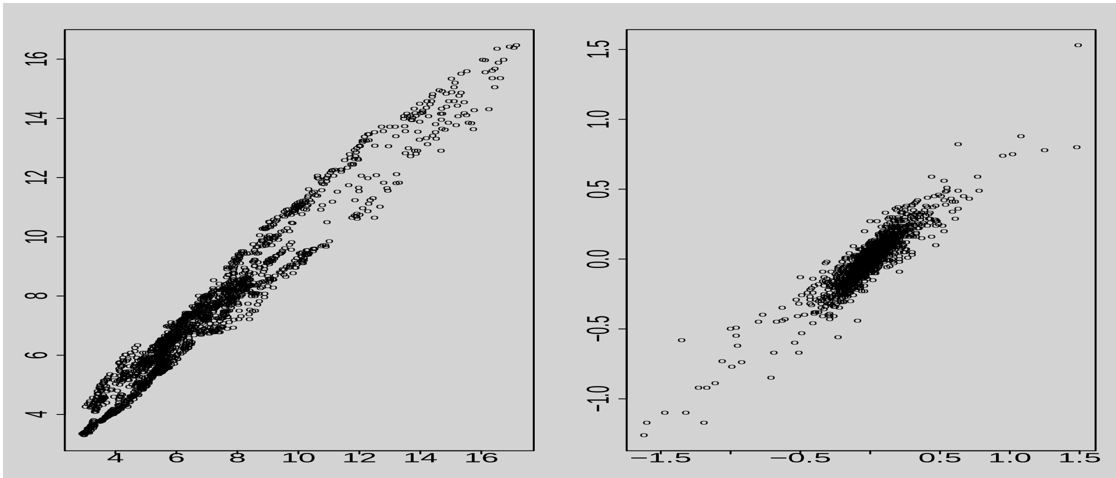

Figure 2.4: Scatterplots of U.S. weekly interest rates from January 5, 1962 to September 10, 1999: the left panel is 3-year rate versus 1-year rate, and the right panel is changes in 3-year rate versus changes in 1-year rate.

and $R^2 = 95.8\%$, where the standard errors of the two coefficients are 0.032 and 0.004, respectively. This simple model (Model I) confirms the high correlation between the two interest rates. However, the model is seriously inadequate as shown by Figure 2.5, which

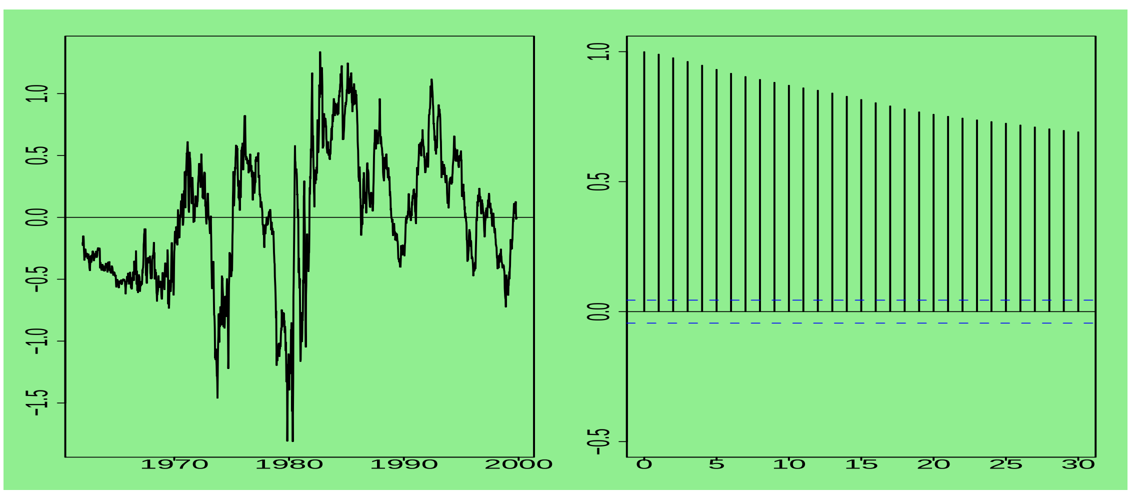

Figure 2.5: Residual series of linear regression Model I for two U.S. weekly interest rates: the left panel is time plot and the right panel is ACF.

gives the time plot and ACF of its residuals. In particular, the sample ACF of the residuals is highly significant and decays slowly, showing the pattern of a **unit root** nonstationary time series[1]. The behavior of the residuals suggests that marked differences exist between the two interest rates. Using the modern econometric terminology, if one assumes that the two

[1] We will discuss in detail on how to do unit root test later

interest rate series are unit root nonstationary, then the behavior of the residuals indicates that the two interest rates are not **co-integrated**; see later chapters for discussion of **unit root** and **co-integration**. In other words, the data fail to support the hypothesis that there exists a long-term equilibrium between the two interest rates. In some sense, this is not surprising because the pattern of "inverted yield curve" did occur during the data span. By the inverted yield curve, we mean the situation under which interest rates are inversely related to their time to maturities.

The unit root behavior of both interest rates and the residuals leads to the consideration of the change series of interest rates. Let $\Delta x_t = y_t - y_{t-1} = (1 - L)\, x_t$ be changes in the 1-year interest rate and $\Delta\, y_t = y_t - y_{t-1} = (1 - L)\, y_t$ denote changes in the 3-year interest rate. Consider the linear regression, Model II: $\Delta\, y_t = \beta_1 + \beta_2\, \Delta\, x_t + e_t$. Figure 2.6 shows time plots of the two change series, whereas the right panel of Figure 2.4 provides a scatterplot

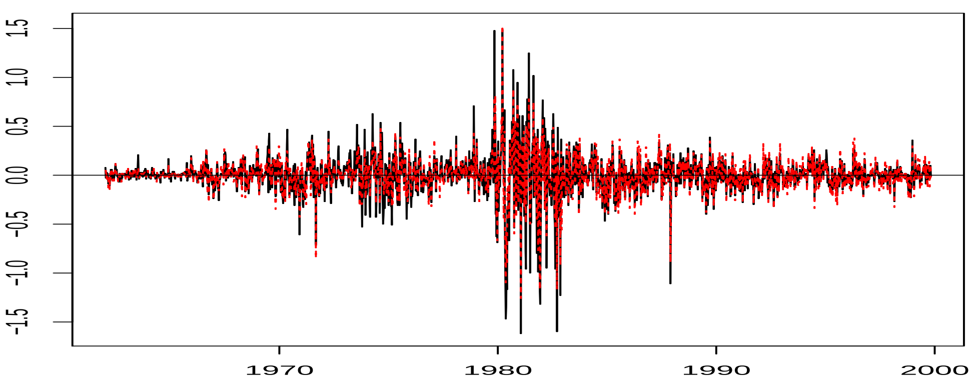

Figure 2.6: Time plots of the change series of U.S. weekly interest rates from January 12, 1962 to September 10, 1999: changes in the Treasury 1-year constant maturity rate are in denoted by black solid line, and changes in the Treasury 3-year constant maturity rate are indicated by red dashed line.

between them. The change series remain highly correlated with a fitted linear regression model given by $\Delta\, y_t = 0.0002 + 0.7811\, \Delta\, x_t + e_t$ with $\widehat{\sigma}_e^2 = 0.0682$ and $R^2 = 84.8\%$. The standard errors of the two coefficients are 0.0015 and 0.0075, respectively. This model further confirms the strong linear dependence between interest rates. The two top panels of Figure 2.7 show the time plot (left) and sample ACF (right) of the residuals (Model II). Once again, the ACF shows some significant serial correlation in the residuals, but the magnitude of the correlation is much smaller. This weak serial dependence in the residuals can be modeled by

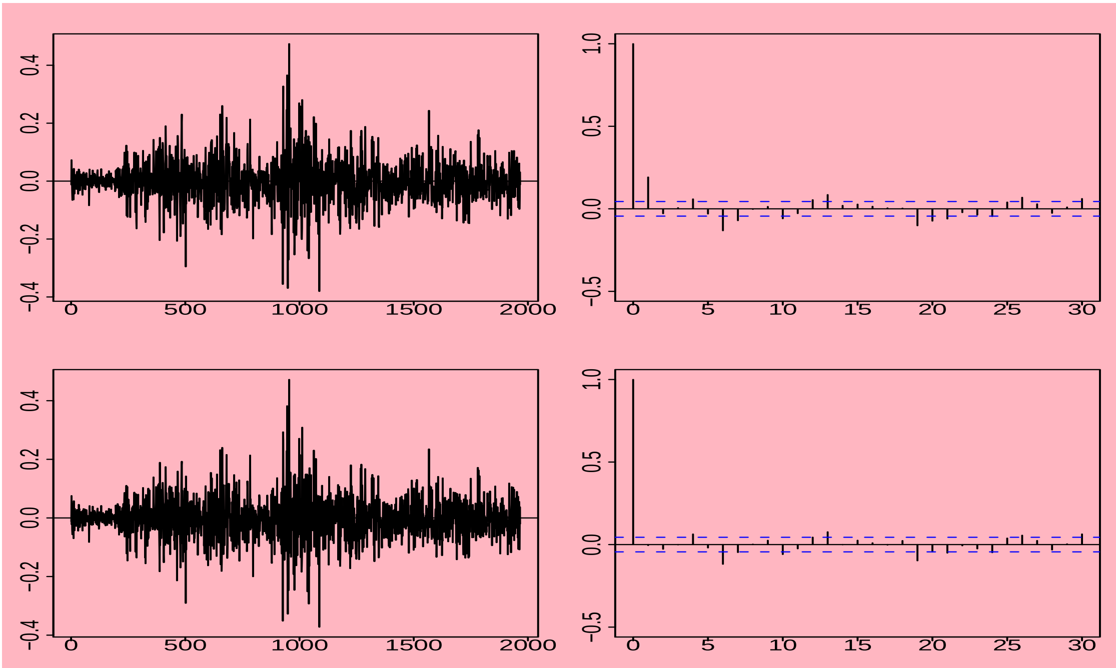

Figure 2.7: Residual series of the linear regression models: Model II (top) and Model III (bottom) for two change series of U.S. weekly interest rates: time plot (left) and ACF (right).

using the simple time series models discussed in the previous sections, and we have a linear regression with time series errors.

The main objective of this section is to discuss a simple approach for building a linear regression model with time series errors. The approach is straightforward. We employ a simple time series model discussed in this chapter for the residual series and estimate the whole model jointly. For illustration, consider the simple linear regression in Model II. Because residuals of the model are serially correlated, we identify a simple ARMA model for the residuals. From the sample ACF of the residuals shown in the right top panel of Figure 2.7, we specify an MA(1) model for the residuals and modify the linear regression model to (Model III): $\Delta\, y_t = \beta_1 + \beta_2\, \Delta\, x_t + e_t$ and $e_t = w_t - \theta_1\, w_{t-1}$, where $\{w_t\}$ is assumed to be a white noise series. In other words, we simply use an MA(1) model, without the constant term, to capture the serial dependence in the error term of Model II. The two bottom panels of Figure 2.7 show the time plot (left) and sample ACF (right) of the residuals (Model III). The resulting model is a simple example of linear regression with time series errors. In practice, more elaborated time series models can be added to a linear regression equation to form a general regression model with time series errors.

Estimating a regression model with time series errors was not easy before the advent of modern computers. Special methods such as the Cochrane-Orcutt estimator have been proposed to handle the serial dependence in the residuals. By now, the estimation is as easy as that of other time series models. If the time series model used is stationary and invertible, then one can estimate the model jointly via the maximum likelihood method or conditional maximum likelihood method. For the U.S. weekly interest rate data, the fitted version of Model II is $\Delta\, y_t = 0.0002 + 0.7824\, \Delta\, x_t + e_t$ and $e_t = w_t + 0.2115\, w_{t-1}$ with $\widehat{\sigma}_w^2 = 0.0668$ and $R^2 = 85.4\%$. The standard errors of the parameters are 0.0018, 0.0077, and 0.0221, respectively. The model no longer has a significant lag-1 residual ACF, even though some minor residual serial correlations remain at lags 4 and 6. The incremental improvement of adding additional MA parameters at lags 4 and 6 to the residual equation is small and the result is not reported here.

Comparing the above three models, we make the following observations. First, the high R^2 and coefficient 0.924 of Modle I are misleading because the residuals of the model show strong serial correlations. Second, for the change series, R^2 and the coefficient of $\Delta\, x_t$ of Model II and Model III are close. In this particular instance, adding the MA(1) model to the change series only provides a marginal improvement. This is not surprising because the estimated MA coefficient is small numerically, even though it is statistically highly significant. Third, the analysis demonstrates that it is important to check residual serial dependence in linear regression analysis. Because the constant term of Model III is insignificant, the model shows that the two weekly interest rate series are related as $y_t = y_{t-1} + 0.782\,(x_t - x_{t-1}) + w_t + 0.212\, w_{t-1}$. The interest rates are concurrently and serially correlated.

Finally, we outline a general procedure for analyzing linear regression models with time series errors: First, fit the linear regression model and check serial correlations of the residuals. Second, if the residual series is unit-root nonstationary, take the first difference of both the dependent and explanatory variables. Go to step 1. If the residual series appears to be stationary, identify an ARMA model for the residuals and modify the linear regression model accordingly. Third, perform a joint estimation via the maximum likelihood method and check the fitted model for further improvement.

To check the serial correlations of residuals, we recommend that the Ljung-Box statistics be used instead of the Durbin-Watson (DW) statistic because the latter only considers the lag-1 serial correlation. There are cases in which residual serial dependence appears at higher order lags. This is particularly so when the time series involved exhibits some seasonal behavior.

Remark: For a residual series e_t with T observations, the Durbin-Watson statistic is

$$DW = \sum_{t=2}^{T} (e_t - e_{t-1})^2 / \sum_{t=1}^{T} e_t^2.$$

Straightforward calculation shows that $DW \approx 2(1 - \widehat{\rho}_e(1))$, where $\rho_e(1)$ is the lag-1 ACF of $\{e_t\}$.

The function in **R** for the Ljung-Box test is

```
Box.test(x, lag = 1, type = c("Box-Pierce", "Ljung-Box"))
```

and the Durbin-Watson test for autocorrelation of disturbances is

```
dwtest(formula, order.by = NULL, alternative = c("greater","two.sided",
 "less"),iterations = 15, exact = NULL, tol = 1e-10, data = list())
```

2.2 Nonparametric Models with Correlated Errors

See the paper by Xiao, Linton, Carrol and Mammen (2003).

2.3 Computer Codes

```
##################################################
# This is Example 3.1 for Johnson and Johnson data
##################################################

y=read.table('c:/res-teach/xiamen12-06/data/ex3-1.dat',header=F)
n=length(y[,1])
y_log=log(y[,1])                            # log of data
postscript(file="c:\\res-teach\\xiamen12-06\\figs\\fig-3.1.eps",
 horizontal=F,width=6,height=6)
```

```
par(mfrow=c(1,2),mex=0.4,bg="light yellow")
ts.plot(y,type="l",lty=1,ylab="",xlab="")
title(main="J&J Earnings",cex=0.5)
ts.plot(y_log,type="l",lty=1,ylab="",xlab="")
title(main="transformed log(earnings)",cex=0.5)
dev.off()

# MODEL 1: y_t=beta_0+beta_1 t+ e_t
z1=1:n
fit1=lm(y_log~z1)                   # fit log(z) versus time trend
e1=fit1$resid
# Now, we need to re-fit the model using the transformed data
x1=5:n
y_1=y_log[5:n]
y_2=y_log[1:(n-4)]
y_fit=y_1-0.7614*y_2
x2=x1-0.7614*(x1-4)
x1=(1-0.7614)*rep(1,n-4)
fit2=lm(y_fit~-1+x1+x2)
e2=fit2$resid
postscript(file="c:\\res-teach\\xiamen12-06\\figs\\fig-3.2.eps",
horizontal=F,width=6,height=6)
par(mfrow=c(2,2),mex=0.4,bg="light pink")
acf(e1, ylab="", xlab="",ylim=c(-0.5,1),lag=30,main="ACF")
text(10,0.8,"detrended")
pacf(e1,ylab="",xlab="",ylim=c(-0.5,1),lag=30,main="PACF")
acf(e2, ylab="", xlab="",ylim=c(-0.5,1),lag=30,main="")
text(15,0.8,"ARIMA(1,0,0,)_4")
pacf(e2,ylab="",xlab="",ylim=c(-0.5,1),lag=30,main="")
dev.off()

#####################################################
```

```
# This is Example 3.2 for weekly interest rate series
#####################################################

z<-read.table("c:/res-teach/xiamen12-06/data/ex3-2.txt",header=F)
#  first column=one year Treasury constant maturity rate;
# second column=three year Treasury constant maturity rate;
#  third column=date

x=z[,1]
y=z[,2]
n=length(x)
u=seq(1962+1/52,by=1/52,length=n)
x_diff=diff(x)
y_diff=diff(y)
# Fit a simple regression model and examine the residuals
fit1=lm(y~x)             # Model 1
e1=fit1$resid

postscript(file="c:\\res-teach\\xiamen12-06\\figs\\fig-3.3.eps",
horizontal=F,width=6,height=6)
matplot(u,cbind(x,y),type="l",lty=c(1,2),col=c(1,2),ylab="",xlab="")
dev.off()

postscript(file="c:\\res-teach\\xiamen12-06\\figs\\fig-3.4.eps",
horizontal=F,width=6,height=6)
par(mfrow=c(1,2),mex=0.4,bg="light grey")
plot(x,y,type="p",pch="o",ylab="",xlab="",cex=0.5)
plot(x_diff,y_diff,type="p",pch="o",ylab="",xlab="",cex=0.5)
dev.off()

postscript(file="c:\\res-teach\\xiamen12-06\\figs\\fig-3.5.eps",
horizontal=F,width=6,height=6)
```

```
par(mfrow=c(1,2),mex=0.4,bg="light green")
plot(u,e1,type="l",lty=1,ylab="",xlab="")
abline(0,0)
acf(e1,ylab="",xlab="",ylim=c(-0.5,1),lag=30,main="")
dev.off()

# Take different and fit a simple regression again
fit2=lm(y_diff~x_diff)            # Model 2
e2=fit2$resid

postscript(file="c:\\res-teach\\xiamen12-06\\figs\\fig-3.6.eps",
horizontal=F,width=6,height=6)
matplot(u[-1],cbind(x_diff,y_diff),type="l",lty=c(1,2),col=c(1,2),
 ylab="",xlab="")
abline(0,0)
dev.off()

postscript(file="c:\\res-teach\\xiamen12-06\\figs\\fig-3.7.eps",
horizontal=F,width=6,height=6)
par(mfrow=c(2,2),mex=0.4,bg="light pink")
ts.plot(e2,type="l",lty=1,ylab="",xlab="")
abline(0,0)
acf(e2, ylab="", xlab="",ylim=c(-0.5,1),lag=30,main="")
# fit a model to the differenced data with an MA(1) error
fit3=arima(y_diff,xreg=x_diff, order=c(0,0,1))      # Model 3
e3=fit3$resid
ts.plot(e3,type="l",lty=1,ylab="",xlab="")
abline(0,0)
acf(e3, ylab="",xlab="",ylim=c(-0.5,1),lag=30,main="")
dev.off()
```

2.4 References

Cochrane, D. and G.H. Orcutt (1949). Applications of least squares regression to relationships containing autocorrelated errors. *Journal of the American Statistical Association*, **44**, 32-61.

Xiao, X., O.B. Linton, R.J. Carroll and E. Mammen (2003). More efficient local polynomial estimation in nonparametric regression with autocorrelated errors. *Journal of the American Statistical Association*, **98**, 480-992.

Chapter 3

Seasonal Time Series Models

3.1 Characteristics of Seasonality

When time series (particularly, economic and financial time series) are observed each day or month or quarter, it is often the case that such as a series displays a seasonal pattern (deterministic cyclical behavior). Similar to the feature of trend, there is no precise definition of **seasonality**. Usually we refer to seasonality when observations in certain seasons display strikingly different features to other seasons. For example, when the retail sales are always large in the fourth quarter (because of the Christmas spending) and small in the first quarter as can be observed from Figure 3.1. It may also be possible that seasonality is reflected in the variance of a time series. For example, for daily observed stock market returns the volatility seems often highest on Mondays, basically because investors have to digest three days of news instead of only day. For mode details, see the book by Taylor (2005, §4.5).

Example 5.1: For Example 3.1, the data shown in Figure 2.1 represent quarterly earnings per share for the American Company Johnson & Johnson from the from the fourth quarter of 1970 to the first quarter of 1980. It is easy to note some very nonstationary behavior in this series that cannot be eliminated completely by differencing or detrending because of the larger fluctuations that occur near the end of the record when the earnings are higher. The right panel of Figure 2.1 shows the log-transformed series and we note that the latter peaks have been attenuated so that the variance of the transformed series seems more stable. One would have to eliminate the trend still remaining in the above series to obtain stationarity. For more details on the current analyses of this series, see the later analyses and the papers by Burman and Shumway (1998) and Cai and Chen (2006).

Example 5.2: In this example we consider the monthly US retail sales series (not seasonally adjusted) from January of 1967 to December of 2000 (in billions of US dollars). The data can be downloaded from the web site at *http://marketvector.com*. The U.S. retail sales index

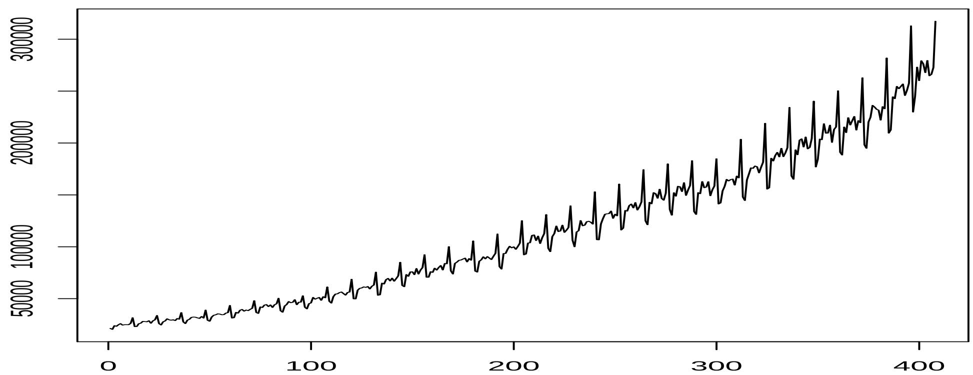

Figure 3.1: US Retail Sales Data from 1967-2000.

is one of the most important indicators of the US economy. There are vast studies of the seasonal series (like this series) in the literature; see, e.g., Franses (1996, 1998) and Ghysels and Osborn (2001) and Cai and Chen (2006). From Figure 3.1, we can observe that the peaks occur in December and we can say that retail sales display seasonality. Also, it can be observed that the trend is basically increasing but nonlinearly. The same phenomenon can be observed from Figure 2.1 for the quarterly earnings for Johnson & Johnson.

If simple graphs are not informative enough to highlight possible seasonal variation, a formal regression model can be used, for example, one might try to consider the following regression model with seasonal dummy variables

$$\Delta\, y_t = y_t - y_{t-1} = \sum_{j=1}^{s} \beta_j\, D_{j,t} + \varepsilon_t,$$

where $D_{j,t}$ is a seasonal dummy variable and s is the number of seasons. Of course, one can use a **seasonal ARIMA** model, denoted by ARIMA$(p, d, q) \times (P, D, Q)_s$, which will be discussed later.

Example 5.3: In this example, we consider a time series with pronounced seasonality displayed in Figure 3.2, where logs of four-weekly advertising expenditures on ratio and television in The Netherlands for $1978.01 - 1994.13$. For these two marketing time series one

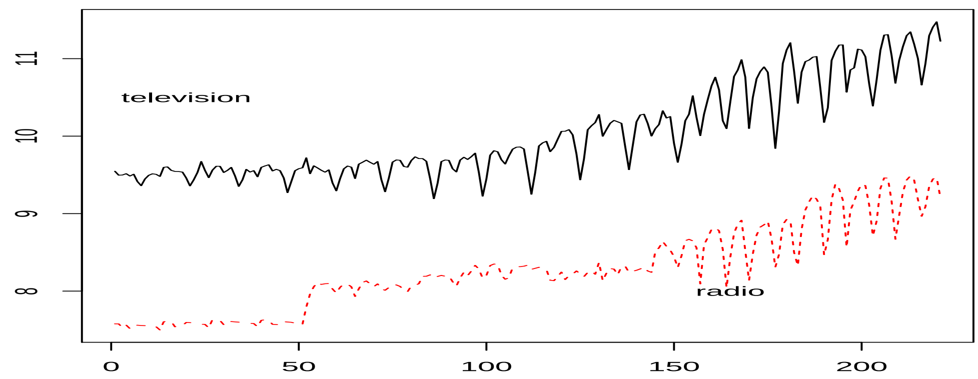

Figure 3.2: Four-weekly advertising expenditures on radio and television in The Netherlands, 1978.01 − 1994.13.

can observe clearly that the television advertising displays quite some seasonal fluctuation throughout the entire sample and the radio advertising has seasonality only for the last five years. Also, there seems to be a structural break in the radio series around observation 53. This break is related to an increase in radio broadcasting minutes in January 1982. Furthermore, there is a visual evidence that the trend changes over time.

Generally, it appears that many time series seasonally observed from business and economics as well as other applied fields display seasonality in the sense that the observations in certain seasons have properties that differ from those data points in other seasons. A second feature of many seasonal time series is that the seasonality changes over time, like what studied by Cai and Chen (2006). Sometimes, these changes appear abrupt, as is the case for advertising on the radio in Figure 3.2, and sometimes such changes occur only slowly. To capture these phenomena, Cai and Chen (2006) proposed a more general flexible seasonal effect model having the following form:

$$y_{ij} = \alpha(t_i) + \beta_j(t_i) + e_{ij}, \qquad i = 1, \ldots, n, \quad j = 1, \ldots, s,$$

where $y_{ij} = y_{(i-1)s+j}$, $t_i = i/n$, $\alpha(\cdot)$ is a (smooth) common trend function in $[0, 1]$, $\{\beta_j(\cdot)\}$ are (smooth) seasonal effect functions in $[0, 1]$, either fixed or random, subject to a set of constraints, and the error term e_{ij} is assumed to be stationary. For more details, see Cai and Chen (2006).

3.2 Modeling

Some economic and financial as well as environmental time series such as quarterly earning per share of a company exhibits certain **cyclical** or **periodic** behavior; see the later chapters on more discussions on **cycles and periodicity**. Such a time series is called a **seasonal** (deterministic cycle) time series. Figure 2.1 shows the time plot of quarterly earning per share of Johnson and Johnson from the first quarter of 1960 to the last quarter of 1980. The data possess some special characteristics. In particular, the earning grew exponentially during the sample period and had a strong seasonality. Furthermore, the variability of earning increased over time. The cyclical pattern repeats itself every year so that the periodicity of the series is 4. If monthly data are considered (e.g., monthly sales of Wal-Mart Stores), then the periodicity is 12. Seasonal time series models are also useful in pricing weather-related derivatives and energy futures.

Analysis of seasonal time series has a long history. In some applications, seasonality is of secondary importance and is removed from the data, resulting in a seasonally adjusted time series that is then used to make inference. The procedure to remove seasonality from a time series is referred to as **seasonal adjustment**. Most economic data published by the U.S. government are seasonally adjusted (e.g., the growth rate of domestic gross product and the unemployment rate). In other applications such as forecasting, seasonality is as important as other characteristics of the data and must be handled accordingly. Because forecasting is a major objective of economic and financial time series analysis, we focus on the latter approach and discuss some econometric models that are useful in modeling seasonal time series.

When the autoregressive, differencing, or seasonal moving average behavior seems to occur at multiples of some underlying period s, a seasonal ARIMA series may result. The seasonal nonstationarity is characterized by slow decay at multiples of s and can often be eliminated by a **seasonal differencing operator** of the form $\Delta_s^D\, x_t = (1 - L^s)^D\, x_t$. For example, when we have monthly data, it is reasonable that a yearly phenomenon will induce $s = 12$ and the ACF will be characterized by slowly decaying spikes at 12, 24, 36, 48, $\cdots$, and we can obtain a stationary series by transforming with the operator $(1 - L^{12})\, x_t = x_t - x_{t-12}$ which is the difference between the current month and the value one year or 12 months ago. If the autoregressive or moving average behavior is seasonal at period s, we define formally the operators

$$\Phi(L^s) = 1 - \Phi_1 L^s - \Phi_2 L^{2s} - \cdots - \Phi_P L^{Ps} \tag{3.1}$$

and

$$\Theta(L^s) = 1 - \Theta_1 L^s - \Theta_2 L^{2s} - \cdots - \Theta_Q L^{Qs}. \tag{3.2}$$

The final form of the seasonal ARIMA$(p, d, q) \times (P, D, Q)_s$ model is

$$\Phi(L^s)\,\phi(L)\Delta_s^D \Delta^d x_t = \Theta(L^s)\,\theta(L)\,w_t. \tag{3.3}$$

Note that one special model of (3.3) is ARIMA$(0, 1, 1) \times (0, 1, 1)_s$, that is

$$(1 - L^s)(1 - L)\,x_t = (1 - \theta_1 L)(1 - \Theta_1 L^s)\,w_t.$$

This model is referred to as the **airline model** or **multiplicative seasonal model** in the literature; see Box and Jenkins (1970), Box, Jenkins, and Reinsel (1994, Chapter 9), and Brockwell and Davis (1991). It has been found to be widely applicable in modeling seasonal time series. The AR part of the model simply consists of the regular and seasonal differences, whereas the MA part involves two parameters.

We may also note the properties below corresponding to **Properties 5.1 - 5.3**.

Property 5.1: *The ACF of a seasonally non-stationary time series decays very slowly at lag multiples s, $2s$, $3s$, $\cdots$, with zeros in between, where s denotes a seasonal period, usually 4 for quarterly data or 12 for monthly data. The PACF of a non-stationary time series tends to have a peak very near unity at lag s.*

Property 5.2: *For a seasonal autoregressive series of order P, the partial autocorrelation function Φ_{hh} as a function of lag h has nonzero values at s, $2s$, $3s$, $\cdots$, Ps, with zeros in between, and is zero for $h > Ps$, the order of the seasonal autoregressive process. There should be some exponential decay.*

Property 5.3: *For a seasonal moving average series of order Q, note that the autocorrelation function (ACF) has nonzero values at s, $2s$, $3s$, $\cdots$, Qs and is zero for $h > Qs$.*

Remark: Note that there is a build-in command in **R** called *arima()* which is a powerful tool for estimating and making inference for an ARIMA model. The command is

```
arima(x,order=c(0,0,0),seasonal=list(order=c(0,0,0),period=NA),
 xreg=NULL,include.mean=TRUE, transform.pars=TRUE,fixed=NULL,init=NULL,
 method=c("CSS-ML","ML","CSS"),n.cond,optim.control=list(),kappa=1e6)
```

See the manuals of **R** for details about this commend.

Example 5.4: We illustrate by fitting the monthly birth series from 1948-1979 shown in Figure 3.3. The period encompasses the boom that followed the Second World War and there

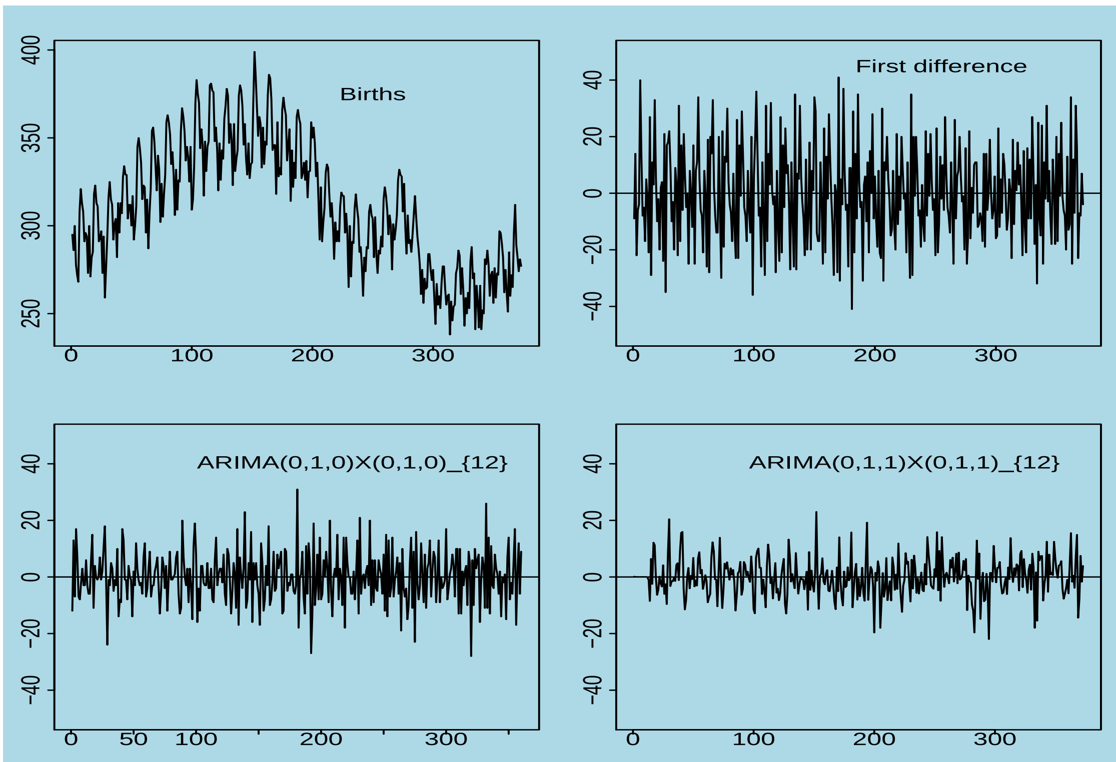

Figure 3.3: Number of live births 1948(1) − 1979(1) and residuals from models with a first difference, a first difference and a seasonal difference of order 12 and a fitted ARIMA$(0,1,1)\times(0,1,1)_{12}$ model.

is the expected rise which persists for about 13 years followed by a decline to around 1974. The series appears to have long-term swings, with seasonal effects superimposed. The long-term swings indicate possible non-stationarity and we verify that this is the case by checking the ACF and PACF shown in the top panel of Figure 3.4. Note that by **Property 5.1**, slow decay of the ACF indicates non-stationarity and we respond by taking a first difference. The results shown in the second panel of Figure 2.5 indicate that the first difference has eliminated the strong low frequency swing. The ACF, shown in the second panel from the top in Figure 3.4 shows peaks at 12, 24, 36, 48, $\cdots$, with now decay. This behavior implies

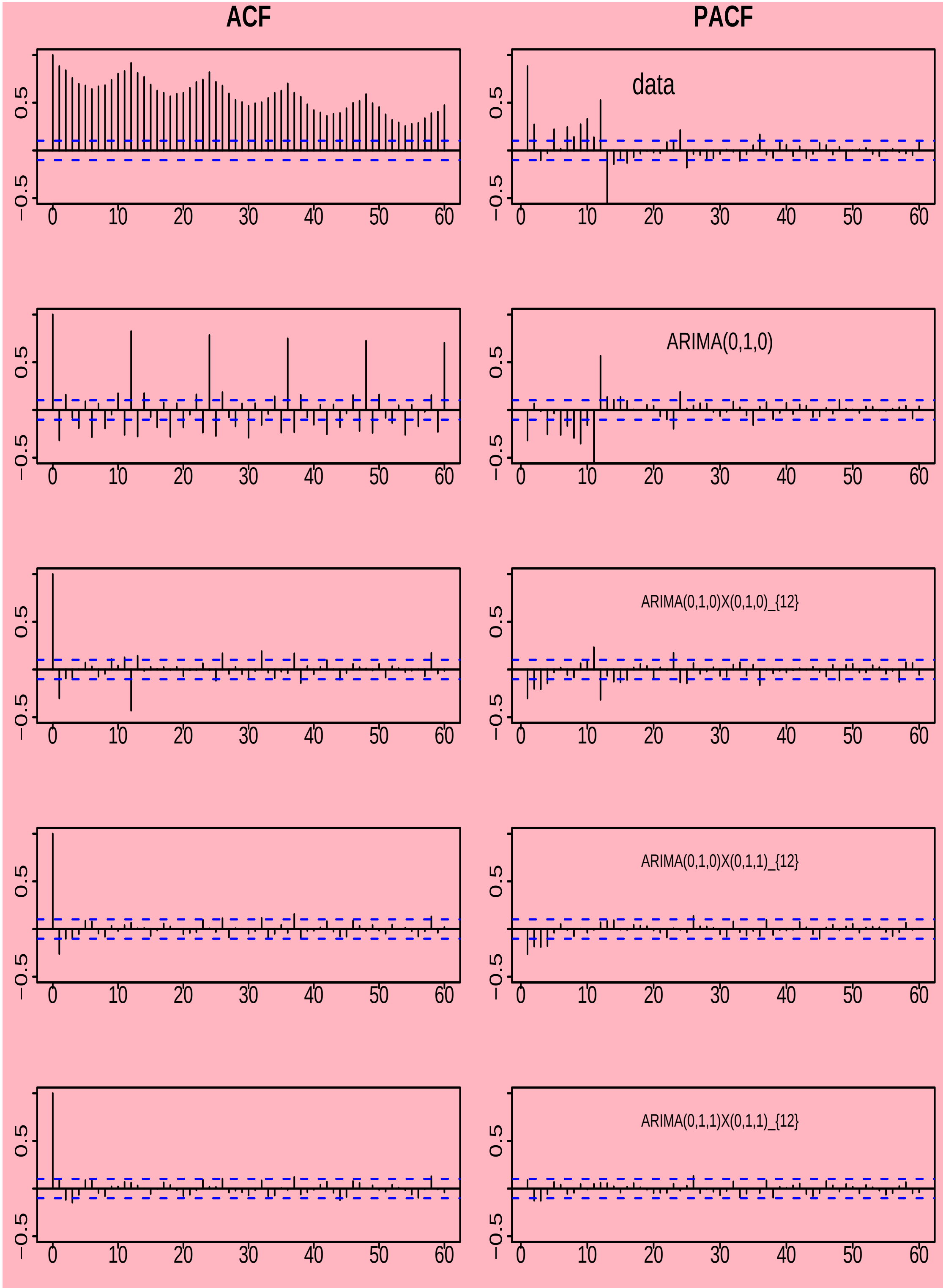

Figure 3.4: Autocorrelation functions and partial autocorrelation functions for the birth series (top two panels), the first difference (second two panels) an ARIMA$(0,1,0)\times(0,1,1)_{12}$ model (third two panels) and an ARIMA$(0,1,1)\times(0,1,1)_{12}$ model (last two panels).

seasonal non-stationarity, by **Property 5.1** above, with $s = 12$. A seasonal difference of the first difference generates an ACF and PACF in Figure 3.4 that we expect for stationary series.

Taking the seasonal difference of the first difference gives a series that looks stationary and has an ACF with peaks at 1 and 12 and a PACF with a substantial peak at 12 and lesser peaks at 12, 24, $\cdots$. This suggests trying either a first order moving average term, or a first order seasonal moving average term with $s = 12$, by **Property 5.3** above. We choose to eliminate the largest peak first by applying a first-order seasonal moving average model with $s = 12$. The ACF and PACF of the residual series from this model, i.e. from ARIMA$(0,1,0)\times(0,1,1)_{12}$, written as $(1-L)(1-L^{12})\,x_t = (1-\Theta_1\,L^{12})\,w_t$, is shown in the fourth panel from the top in Figure 3.4. We note that the peak at lag one is still there, with attending exponential decay in the PACF. This can be eliminated by fitting a first-order moving average term and we consider the model ARIMA$(0,1,1)\times(0,1,1)_{12}$, written as

$$(1-L)(1-L^{12})\,x_t = (1-\theta_1\,L)(1-\Theta_1\,L^{12})\,w_t.$$

The ACF of the residuals from this model are relatively well behaved with a number of peaks either near or exceeding the 95% test of no correlation. Fitting this final ARIMA$(0,1,1)\times(0,1,1)_{12}$ model leads to the model

$$(1-L)(1-L^{12})\,x_t = (1-0.4896\,L)(1-0.6844\,L^{12})\,w_t$$

with AICC$= 4.95$, $R^2 = 0.9804^2 = 0.961$, and the p-values are $(0.000,\ 0.000)$. The ARIMA search leads to the model

$$(1-L)(1-L^{12})\,x_t = (1-0.4088\,L-0.1645\,L^2)(1-0.6990\,L^{12})\,w_t,$$

yielding AICC$= 4.92$ and $R^2 = 0.981^2 = 0.962$, slightly better than the ARIMA$(0,1,1)\times(0,1,1)_{12}$ model. Evaluating these latter models leads to the conclusion that the extra parameters do not add a practically substantial amount to the predictability. The model is expanded as

$$x_t = x_{t-1} + x_{t-12} - x_{t-13} + w_t - \theta_1\,w_{t-1} - \Theta_1\,w_{t-12} + \theta_1\Theta_1\,w_{t-13}.$$

The forecast is

$$x_{t+1}^t = x_t + x_{t-11} - x_{t-12} - \theta_1\,w_t - \Theta_1\,w_{t-11} + \theta_1\,\Theta_1\,w_{t-12}$$

$$x_{t+2}^t = x_{t+1}^t + x_{t-10} - x_{t-11} - \Theta_1\, w_{t-10} + \theta_1 \Theta_1\, w_{t-11}.$$

Continuing in the same manner, we obtain

$$x_{t+12}^t = x_{t+11}^t + x_t - x_{t-1} - \Theta_1\, w_t + \theta_1 \Theta_1\, w_{t-1}$$

for the 12 month forecast.

Example 5.5: Figure 3.5 shows the autocorrelation function of the log-transformed J&J earnings series that is plotted in Figure 2.1 and we note the slow decay indicating the nonstationarity which has already been obvious in the Chapter 3 discussion. We may also compare the ACF with that of a random walk, and note the close similarity. The partial autocorrelation function is very high at lag one which, under ordinary circumstances, would indicate a first order autoregressive AR(1) model, except that, in this case, the value is close to unity, indicating a root close to 1 on the unit circle. The only question would be whether differencing or detrending is the better transformation to stationarity. Following in the Box-Jenkins tradition, differencing leads to the ACF and PACF shown in the second panel and no simple structure is apparent. To force a next step, we interpret the peaks at 4, 8, 12, 16, $\cdots$, as contributing to a possible seasonal autoregressive term, leading to a possible ARIMA$(0,1,0) \times (1,0,0)_4$ and we simply fit this model and look at the ACF and PACF of the residuals, shown in the third two panels. The fit improves somewhat, with significant peaks still remaining at lag 1 in both the ACF and PACF. The peak in the ACF seems more isolated and there remains some exponentially decaying behavior in the PACF, so we try a model with a first-order moving average. The bottom two panels show the ACF and PACF of the resulting ARIMA$(0,1,1)\times(1,0,0)_4$ and we note only relatively minor excursions above and below the 95% intervals under the assumption that the theoretical ACF is white noise. The final model suggested is ($y_t = \log x_t$)

$$(1 - \Phi_1\, L^4)(1 - L)\, y_t = (1 - \theta_1\, L)\, w_t, \tag{3.4}$$

where $\widehat{\Phi}_1 = 0.820(0.058)$, $\widehat{\theta}_1 = 0.508(0.098)$, and $\widehat{\sigma}_w^2 = 0.0086$. The model can be written in forecast form as

$$y_t = y_{t-1} + \Phi_1(y_{t-4} - y_{t-5}) + w_t - \theta_1\, w_{t-1}.$$

The residual plot of the above is plotted in the left bottom panel of Figure 3.6. To forecast the original series for, say 4 quarters, we compute the forecast limits for $y_t = \log x_t$ and then exponentiate, i.e. $x_{t+h}^t = \exp(y_{t+h}^t)$.

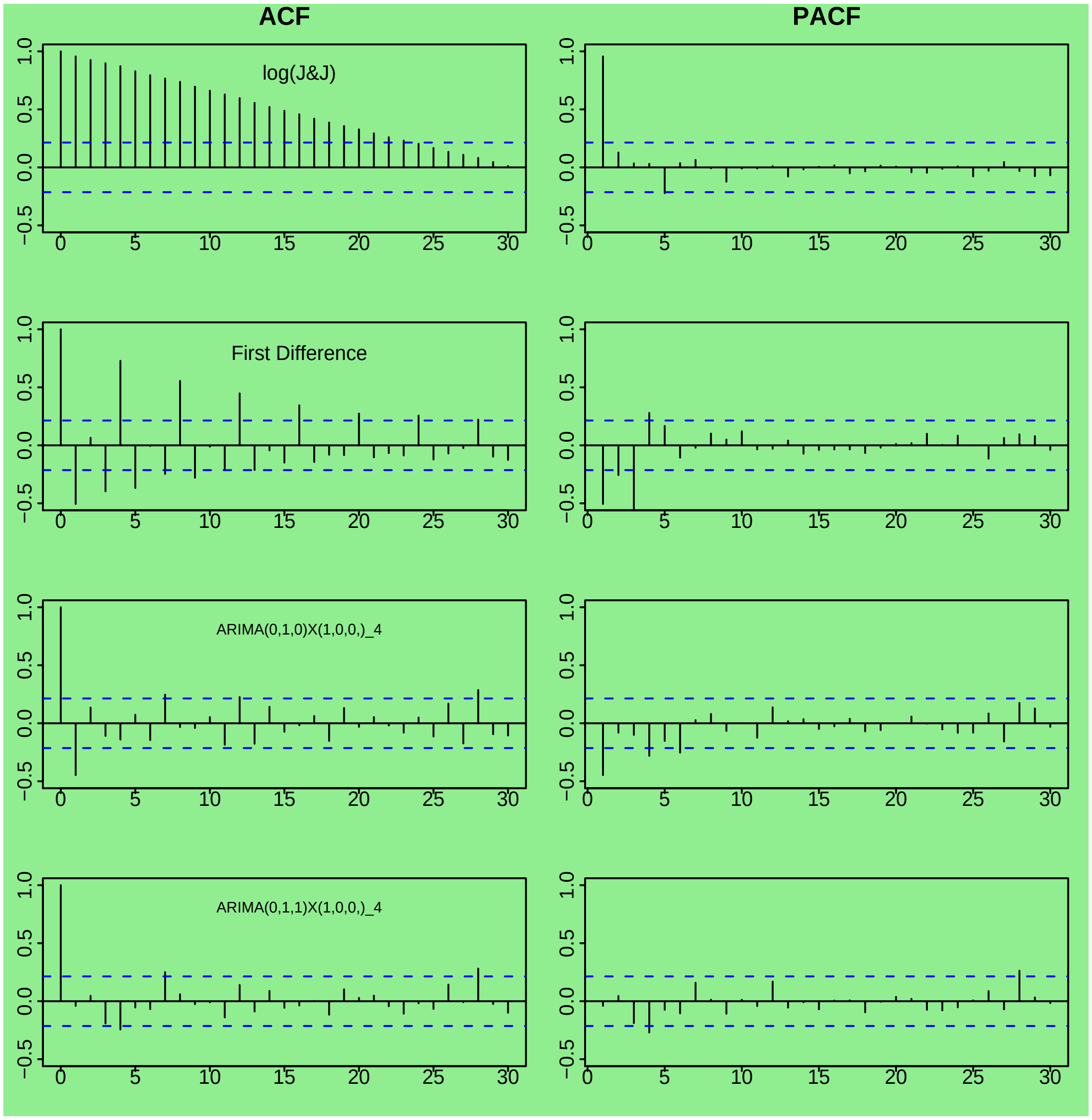

Figure 3.5: Autocorrelation functions (ACF) and partial autocorrelation functions (PACF) for the log J&J earnings series (top two panels), the first difference (second two panels), ARIMA$(0,1,0)\times(1,0,0)_4$ model (third two panels), and ARIMA$(0,1,1)\times(1,0,0)_4$ model (last two panels).

Based on the the exact likelihood method, Tsay (2005) considered the following seasonal ARIMA$(0,1,1)\times(0,1,1)_4$ model

$$(1-L)(1-L^4)\,y_t=(1-0.678\,L)(1-0.314\,L^4)\,w_t, \tag{3.5}$$

with $\widehat{\sigma}_w^2=0.089$, where standard errors of the two MA parameters are 0.080 and 0.101, respectively. The Ljung-Box statistics of the residuals show $Q(12)=10.0$ with p-value 0.44. The model appears to be adequate. The ACF and PACF of the ARIMA$(0,1,1)\times(0,1,1)_4$

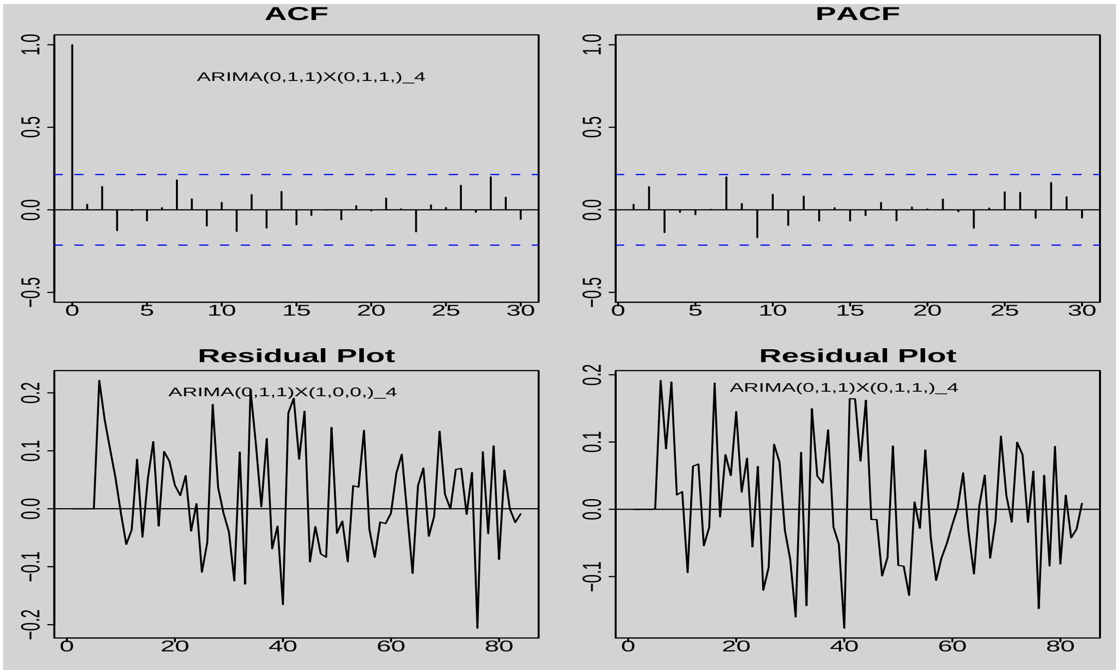

Figure 3.6: ACF and PACF for ARIMA$(0, 1, 1) \times (0, 1, 1)_4$ model (top two panels) and the residual plots of ARIMA$(0, 1, 1) \times (1, 0, 0)_4$ (left bottom panel) and ARIMA$(0, 1, 1) \times (0, 1, 1)_4$ model (right bottom panel).

model are given in the top two panels of Figure 3.6 and the residual plot is displayed in the right bottom panel of Figure 3.6. Based on the comparison of ACF and PACF of two model (3.4) and (3.5) [the last two panels of Figure 3.5 and the top two panels in Figure 3.6], it seems that ARIMA$(0, 1, 1) \times (0, 1, 1)_4$ model in (3.5) might perform better than ARIMA$(0, 1, 1) \times (1, 0, 0)_4$ model in (3.4).

To illustrate the forecasting performance of the seasonal model in (3.5), we re-estimate the model using the first 76 observations and reserve the last eight data points for forecasting evaluation. We compute 1-step to 8-step ahead forecasts and their standard errors of the fitted model at the forecast origin $t = 76$. An anti-log transformation is taken to obtain forecasts of earning per share using the relationship between normal and log-normal distributions. Figure 2.15 in Tsay (2005, p.77) shows the forecast performance of the model, where the observed data are in solid line, point forecasts are shown by dots, and the dashed lines show 95% interval forecasts. The forecasts show a strong seasonal pattern and are close to the observed data. For more comparisons for forecasts using different models including semiparametric and nonparametric models, the reader is referred to the book by Shumway (1988), and Shumway and Stoffer (2000) and the papers by Burman and Shummay (1998) and Cai and Chen (2006).

When the seasonal pattern of a time series is stable over time (e.g., close to a deterministic function), dummy variables may be used to handle the seasonality. This approach is taken by some analysts. However, deterministic seasonality is a special case of the multiplicative seasonal model discussed before. Specifically, if $\Theta_1 = 1$, then model contains a deterministic seasonal component. Consequently, the same forecasts are obtained by using either dummy variables or a multiplicative seasonal model when the seasonal pattern is deterministic. Yet use of dummy variables can lead to inferior forecasts if the seasonal pattern is not deterministic. In practice, we recommend that the exact likelihood method should be used to estimate a multiplicative seasonal model, especially when the sample size is small or when there is the possibility of having a deterministic seasonal component.

Example 5.6: To determine deterministic behavior, consider the monthly simple return of the CRSP Decile 1 index from January 1960 to December 2003 for 528 observations. The series is shown in the left top panel of Figure 3.7 and the time series does not show any clear pattern of seasonality. However, the sample ACf of the return series shown in the left

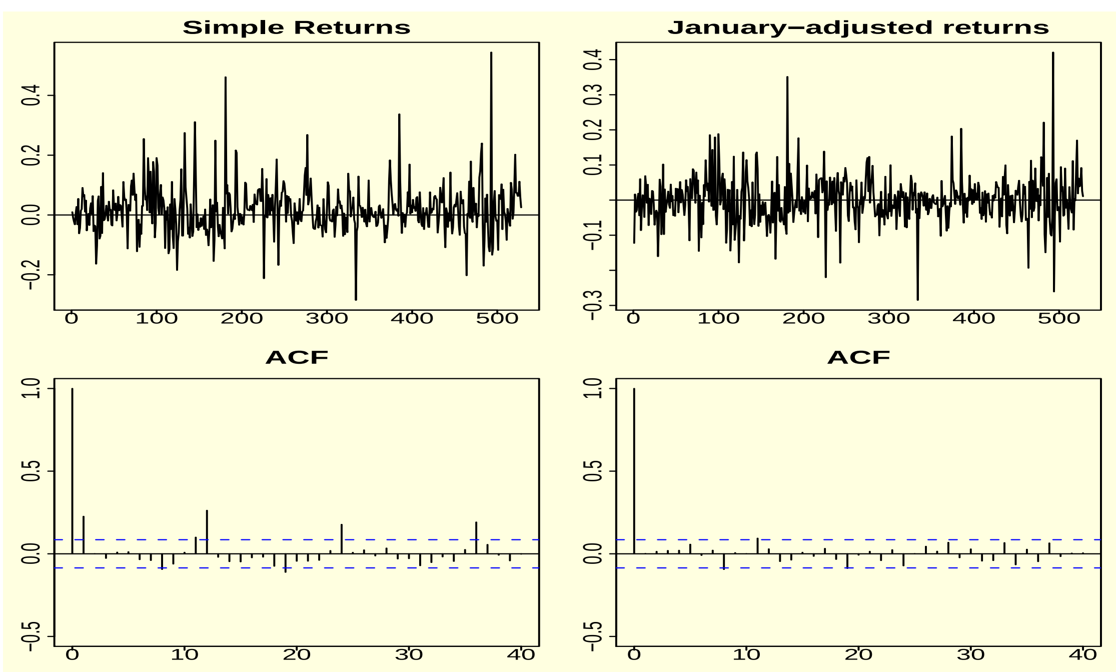

Figure 3.7: Monthly simple return of CRSP Decile 1 index from January 1960 to December 2003: Time series plot of the simple return (left top panel), time series plot of the simple return after adjusting for January effect (right top panel), the ACF of the simple return (left bottom panel), and the ACF of the adjusted simple return.

bottom panel of Figure 3.7 contains significant lags at 12, 24, and 36 as well as lag 1. If seasonal AIMA models are entertained, a model in form

$$(1-\phi_1\,L)(1-\Phi_1\,L^{12})\,x_t=\alpha+(1-\Theta_1\,L^{12})\,w_t$$

is identified, where x_t is the monthly simple return. Using the conditional likelihood, the fitted model is

$$(1-0.25\,L)(1-0.99\,L^{12})\,x_t=0.0004+(1-0.92\,L^{12})\,w_t$$

with $\sigma_w=0.071$. The MA coefficient is close to unity, indicating that the fitted model is close to being non-invertible. If the exact likelihood method is used, we have

$$(1-0.264\,L)(1-0.996\,L^{12})\,x_t=0.0002+(1-0.999\,L^{12})\,w_t$$

with $\sigma_w=0.067$. Cancellation between seasonal AR and MA factors is clearly. This highlights the usefulness of using the exact likelihood method, and the estimation result suggests that the seasonal behavior might be deterministic. To further confirm this assertion, we define the dummy variable for January, that is

$$J_t=\begin{cases}1 & \text{if } t \text{ is January,}\\ 0 & \text{otherwise,}\end{cases}$$

and employ the simple linear regression

$$x_t=\beta_0+\beta_1\,J_t+e_t.$$

The right panels of Figure 3.7 show the time series plot of and the ACF of the residual series of the prior simple linear regression. From the ACF, there are no significant serial correlation at any multiples of 12, suggesting that the seasonal pattern has been successfully removed by the January dummy variable. Consequently, the seasonal behavior in the monthly simple return of Decile 1 is due to the *January effect*.

3.3 Nonlinear Seasonal Time Series Models

See the papers by Burman and Shumway (1998) and Cai and Chen (2006) and the books by Franses (1998) and Ghysels and Osborn (2001). The reading materials are the papers by Burman and Shumway (1998) and Cai and Chen (2006).

3.4 Computer Codes

```
##########################################
# This is Example 5.2 for retail sales data
#########################################
y=read.table("c:/res-teach/xiamen12-06/data/ex5-2.txt",header=F)
postscript(file="c:/res-teach/xiamen12-06/figs/fig-5.1.eps",
 horizontal=F,width=6,height=6)
ts.plot(y,type="l",lty=1,ylab="",xlab="")
dev.off()

############################################
# This is Example 5.3 for the marketing data
############################################
text_tv=c("television")
text_radio=c("radio")
data<-read.table("c:/res-teach/xiamen12-06/data/ex5-3.txt",header=T)
TV=log(data[,1])
RADIO=log(data[,2])
postscript(file="c:/res-teach/xiamen12-06/figs/fig-5.2.eps",
 horizontal=F,width=6,height=6)
ts.plot(cbind(TV,RADIO),type="l",lty=c(1,2),col=c(1,2),ylab="",xlab="")
text(20,10.5,text_tv)
text(165,8,text_radio)
dev.off()

 ####################################################################
# This is Example 5.4 in Chapter 2
####################################
x<-matrix(scan("c:/res-teach/xiamen12-06/data/ex5-4.txt"),byrow=T,ncol=1)
n=length(x)
x_diff=diff(x)
x_diff_12=diff(x_diff,lag=12)
```

```
fit1=arima(x,order=c(0,0,0),seasonal=list(order=c(0,0,0)),include.mean=F)
resid_1=fit1$resid
fit2=arima(x,order=c(0,1,0),seasonal=list(order=c(0,0,0)),include.mean=F)
resid_2=fit2$resid
fit3=arima(x,order=c(0,1,0),seasonal=list(order=c(0,1,0),period=12),
 include.mean=F)
resid_3=fit3$resid

postscript(file="c:/res-teach/xiamen12-06/figs/fig-5.4.eps",
horizontal=F,width=6,height=6)
 par(mfrow=c(5,2),mex=0.4,bg="light pink")
acf(resid_1, ylab="", xlab="",ylim=c(-0.5,1),lag=60,main="ACF",cex=0.7)
pacf(resid_1,ylab="",xlab="",ylim=c(-0.5,1),lag=60,main="PACF",cex=0.7)
text(20,0.7,"data",cex=1.2)
acf(resid_2, ylab="", xlab="",ylim=c(-0.5,1),lag=60,main="")
 # differenced data
pacf(resid_2,ylab="",xlab="",ylim=c(-0.5,1),lag=60,main="")
text(30,0.7,"ARIMA(0,1,0)")
acf(resid_3, ylab="", xlab="",ylim=c(-0.5,1),lag=60,main="")
 # seasonal difference of differenced data
pacf(resid_3,ylab="",xlab="",ylim=c(-0.5,1),lag=60,main="")
text(30,0.7,"ARIMA(0,1,0)X(0,1,0)_{12}",cex=0.8)

fit4=arima(x,order=c(0,1,0),seasonal=list(order=c(0,1,1),
 period=12),include.mean=F)
resid_4=fit4$resid
fit5=arima(x,order=c(0,1,1),seasonal=list(order=c(0,1,1),
 period=12),include.mean=F)
resid_5=fit5$resid

acf(resid_4, ylab="", xlab="",ylim=c(-0.5,1),lag=60,main="")
```

```
 # ARIMA(0,1,0)*(0,1,1)_12
pacf(resid_4,ylab="",xlab="",ylim=c(-0.5,1),lag=60,main="")
text(30,0.7,"ARIMA(0,1,0)X(0,1,1)_{12}",cex=0.8)
acf(resid_5, ylab="", xlab="",ylim=c(-0.5,1),lag=60,main="")
 # ARIMA(0,1,1)*(0,1,1)_12
pacf(resid_5,ylab="",xlab="",ylim=c(-0.5,1),lag=60,main="")
text(30,0.7,"ARIMA(0,1,1)X(0,1,1)_{12}",cex=0.8)
dev.off()

postscript(file="c:/res-teach/xiamen12-06/figs/fig-5.3.eps",
 horizontal=F,width=6,height=6)
par(mfrow=c(2,2),mex=0.4,bg="light blue")
ts.plot(x,type="l",lty=1,ylab="",xlab="")
text(250,375, "Births")
ts.plot(x_diff,type="l",lty=1,ylab="",xlab="",ylim=c(-50,50))
text(255,45, "First difference")
abline(0,0)
ts.plot(x_diff_12,type="l",lty=1,ylab="",xlab="",ylim=c(-50,50))
   # time series plot of the seasonal difference (s=12) of differenced data
text(225,40,"ARIMA(0,1,0)X(0,1,0)_{12}")
abline(0,0)
ts.plot(resid_5,type="l",lty=1,ylab="",xlab="",ylim=c(-50,50))
text(225,40, "ARIMA(0,1,1)X(0,1,1)_{12}")
abline(0,0)
dev.off()

########################
# This is Example 5.5
########################

y=read.table('c:/res-teach/xiamen12-06/data/ex3-1.txt',header=F)
n=length(y[,1])
```

```
y_log=log(y[,1])                            # log of data
y_diff=diff(y_log)            # first-order difference
y_diff_4=diff(y_diff,lag=4)   # first-order seasonal difference

fit1=ar(y_log,order=1)        # fit AR(1) model
#print(fit1)
library(tseries)              # call library(tseries)
library(zoo)
fit1_test=adf.test(y_log)
# do Augmented Dicky-Fuller test for tesing unit root
#print(fit1_test)

fit1=arima(y_log,order=c(0,0,0),seasonal=list(order=c(0,0,0)),
 include.mean=F)
resid_21=fit1$resid
fit2=arima(y_log,order=c(0,1,0),seasonal=list(order=c(0,0,0)),
 include.mean=F)
resid_22=fit2$resid         # residual for ARIMA(0,1,0)*(0,0,0)
fit3=arima(y_log,order=c(0,1,0),seasonal=list(order=c(1,0,0),period=4),
 include.mean=F,method=c("CSS"))
resid_23=fit3$resid         # residual for ARIMA(0,1,0)*(1,0,0)_4
# note that this model is non-stationary so that "CSS" is used

postscript(file="c:\\res-teach\\xiamen12-06\\figs\\fig-5.5.eps",
horizontal=F,width=6,height=6)
par(mfrow=c(4,2),mex=0.4,bg="light green")
acf(resid_21, ylab="", xlab="",ylim=c(-0.5,1),lag=30,main="ACF",cex=0.7)
text(16,0.8,"log(J&J)")
pacf(resid_21,ylab="",xlab="",ylim=c(-0.5,1),lag=30,main="PACF",cex=0.7)
acf(resid_22, ylab="", xlab="",ylim=c(-0.5,1),lag=30,main="")
text(16,0.8,"First Difference")
pacf(resid_22,ylab="",xlab="",ylim=c(-0.5,1),lag=30,main="")
```

```
acf(resid_23, ylab="", xlab="",ylim=c(-0.5,1),lag=30,main="")
text(16,0.8,"ARIMA(0,1,0)X(1,0,0,)_4",cex=0.8)
pacf(resid_23,ylab="",xlab="",ylim=c(-0.5,1),lag=30,main="")
fit4=arima(y_log,order=c(0,1,1),seasonal=list(order=c(1,0,0),
 period=4),include.mean=F,method=c("CSS"))
resid_24=fit4$resid      # residual for ARIMA(0,1,1)*(1,0,0)_4
# note that this model is non-stationary
#print(fit4)
fit4_test=Box.test(resid_24,lag=12, type=c("Ljung-Box"))
#print(fit4_test)
acf(resid_24, ylab="", xlab="",ylim=c(-0.5,1),lag=30,main="")
text(16,0.8,"ARIMA(0,1,1)X(1,0,0,)_4",cex=0.8)
 # ARIMA(0,1,1)*(1,0,0)_4
pacf(resid_24,ylab="",xlab="",ylim=c(-0.5,1),lag=30,main="")
dev.off()

fit5=arima(y_log,order=c(0,1,1),seasonal=list(order=c(0,1,1),period=4),
 include.mean=F,method=c("ML"))
resid_25=fit5$resid       # residual for ARIMA(0,1,1)*(0,1,1)_4
#print(fit5)
fit5_test=Box.test(resid_25,lag=12, type=c("Ljung-Box"))
#print(fit5_test)

postscript(file="c:\\res-teach\\xiamen12-06\\figs\\fig-5.6.eps",
horizontal=F,width=6,height=6,bg="light grey")
par(mfrow=c(2,2),mex=0.4)
acf(resid_25, ylab="", xlab="",ylim=c(-0.5,1),lag=30,main="ACF")
text(16,0.8,"ARIMA(0,1,1)X(0,1,1,)_4",cex=0.8)
           # ARIMA(0,1,1)*(0,1,1)_4
pacf(resid_25,ylab="",xlab="",ylim=c(-0.5,1),lag=30,main="PACF")
ts.plot(resid_24,type="l",lty=1,ylab="",xlab="")
title(main="Residual Plot",cex=0.5)
```

```
  text(40,0.2,"ARIMA(0,1,1)X(1,0,0,)_4",cex=0.8)
abline(0,0)
ts.plot(resid_25,type="l",lty=1,ylab="",xlab="")
title(main="Residual Plot",cex=0.5)
 text(40,0.18,"ARIMA(0,1,1)X(0,1,1,)_4",cex=0.8)
abline(0,0)
dev.off()

#########################
# This is Example 5.6
########################

z<-matrix(scan("c:/res-teach/xiamen12-06/data/ex5-6.txt"),byrow=T,ncol=4)
decile1=z[,2]

# Model 1: an ARIMA(1,0,0)*(1,0,1)_12
fit1=arima(decile1,order=c(1,0,0),seasonal=list(order=c(1,0,1),
 period=12),include.mean=T)
#print(fit1)
e1=fit1$resid
n=length(decile1)
m=n/12
jan=rep(c(1,0,0,0,0,0,0,0,0,0,0,0),m)
feb=rep(c(0,1,0,0,0,0,0,0,0,0,0,0),m)
mar=rep(c(0,0,1,0,0,0,0,0,0,0,0,0),m)
apr=rep(c(0,0,0,1,0,0,0,0,0,0,0,0),m)
may=rep(c(0,0,0,0,1,0,0,0,0,0,0,0),m)
jun=rep(c(0,0,0,0,0,1,0,0,0,0,0,0),m)
jul=rep(c(0,0,0,0,0,0,1,0,0,0,0,0),m)
aug=rep(c(0,0,0,0,0,0,0,1,0,0,0,0),m)
sep=rep(c(0,0,0,0,0,0,0,0,1,0,0,0),m)
oct=rep(c(0,0,0,0,0,0,0,0,0,1,0,0),m)
```

```
nov=rep(c(0,0,0,0,0,0,0,0,0,0,1,0),m)
dec=rep(c(0,0,0,0,0,0,0,0,0,0,0,1),m)
de=cbind(decile1[jan==1],decile1[feb==1],decile1[mar==1],decile1[apr==1],
 decile1[may==1],decile1[jun==1],decile1[jul==1],decile1[aug==1],
 decile1[sep==1],decile1[oct==1],decile1[nov==1],decile1[dec==1])
 # Model 2: a simple regression model without correlated errors
# to see the effect from January
fit2=lm(decile1~jan)
e2=fit2$resid
#print(summary(fit2))
# Model 3: a regression model with correlated errors
fit3=arima(decile1,xreg=jan,order=c(0,0,1),include.mean=T)
e3=fit3$resid
#print(fit3)

postscript(file="c:/res-teach/xiamen12-06/figs/fig-5.7.eps",
horizontal=F,width=6,height=6)
par(mfrow=c(2,2),mex=0.4,bg="light yellow")
ts.plot(decile1,type="l",lty=1,col=1,ylab="",xlab="")
title(main="Simple Returns",cex=0.5)
abline(0,0)
ts.plot(e3,type="l",lty=1,col=1,ylab="",xlab="")
title(main="January-adjusted returns",cex=0.5)
abline(0,0)
acf(decile1, ylab="", xlab="",ylim=c(-0.5,1),lag=40,main="ACF")
acf(e3,ylab="",xlab="",ylim=c(-0.5,1),lag=40,main="ACF")
dev.off()
```

3.5 References

Box, G.E.P. and Jenkins, G.M. (1970). *Time Series Analysis, Forecasting, and Control*. Holden Day, San Francisco.

Box, G.E.P., G.M. Jenkins and G.C. Reinsel (1994). *Time Series Analysis, Forecasting and Control*. 3th Edn. Englewood Cliffs, NJ: Prentice-Hall.

Brockwell, P.J. and Davis, R.A. (1991). *Time Series Theory and Methods*. New York: Springer.

Burman, P. and R.H. Shumway (1998). Semiparametric modeling of seasonal time series. Journal of Time Series Analysis, **19**, 127-145.

Cai, Z. and R. Chen (2006). Flexible seasonal time series models. *Advances in Econometrics*, **20B**, 63-87.

Franses, P.H. (1998). *Time Series Models for Business and Economic Forecasting*. New York: Cambridge University Press.

Franses, P.H. and D. van Dijk (2000). *Nonlinear Time Series Models for Empirical Finance*. New York: Cambridge University Press.

Ghysels, E. and D.R. Osborn (2001). *The Econometric Analysis of Seasonal Time Series*. New York: Cambridge University Press.

Shumway, R.H. (1988). *Applied Statistical Time Series Analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Shumway, R.H., A.S. Azari and Y. Pawitan (1988). Modeling mortality fluctuations in Los Angeles as functions of pollution and weather effects. *Environmental Research*, **45**, 224-241.

Shumway, R.H. and D.S. Stoffer (2000). *Time Series Analysis & Its Applications*. New York: Springer-Verlag.

Tiao, G.C. and R.S. Tsay (1983). Consistency properties of least squares estimates of autoregressive parameters in ARMA models. *Annals of Statistics*, **11**, 856-871.

Tsay, R.S. (2005). *Analysis of Financial Time Series*, 2th Edition. John Wiley & Sons, New York.

Chapter 4

Long Memory Models and Structural Changes

Long memory time series have been a popular area of research in economics, finance and statistics and other applied fields such as hydrological sciences during the recent years. Long memory dependence was first observed by the hydrologist Hurst (1951) when analyzing the minimal water flow of the Nile River when planning the Aswan Dam. Granger (1966) gave an intensive discussion about the application of long memory dependence in economics and its consequence was initiated. But in many applications, it is not clear whether the observed dependence structure is real long memory or an artefact of some other phenomenon such as structural breaks or deterministic trends. Long memory in the data would have strong consequences.

4.1 Long Memory Models

4.1.1 Methodology

We have discussed that for a stationary time series the ACF decays exponentially to zero as lag increases. Yet for a unit root nonstationary time series, it can be shown that the sample ACF converges to 1 for all fixed lags as the sample size increases; see Chan and Wei (1988) and Tiao and Tsay (1983). There exist some time series whose ACF decays slowly to zero at a polynomial rate as the lag increases. These processes are referred to as **long memory** or **long range dependent** time series. One such an example is the **fractionally differenced process** defined by

$$(1-L)^d\, x_t = w_t, \quad |d| < 0.5, \tag{4.1}$$

where $\{w_t\}$ is a white noise series and d is called the long memory parameter or $H = d + 1/2$ is called the Hurst parameter; see Hurst (1951). Properties of model (4.1) have been widely studied in the literature (e.g., Beran, 1994). We summarize some of these properties below.

1. If $d < 0.5$, then x_t is a weakly stationary process and has the infinite MA representation
$$x_t = w_t + \sum_{k=1}^{\infty} \psi_k \, w_{t-k} \quad \text{with} \quad \psi_k = d(d+1)\cdots(d+k-1)/k! = \binom{k+d-1}{k}.$$

2. If $d > -0.5$, then x_t is invertible and has the infinite AR representation.
$$x_t = w_t + \sum_{k=1}^{\infty} \psi_k \, w_{t-k} \quad \text{with} \quad \psi_k = (0-d)(1-d)\cdots(k-1-d)/k! = \binom{k-d-1}{k}.$$

3. For $|d| < 0.5$, the ACF of x_t is
$$\rho_x(h) = \frac{d(1+d)\cdots(h-1+d)}{(1-d)(2-d)\cdots(h-d)}, \qquad h \geq 1.$$
In particular, $\rho_x(1) = d/(1-d)$ and as $h \to \infty$,
$$\rho_x(h) \approx \frac{(-d)!}{(d-1)!} \, h^{2d-1}.$$

4. For $|d| < 0.5$, the PACF of x_t is $\phi_{h,h} = d/(h-d)$ for $h \geq 1$.

5. For $|d| < 0.5$, the spectral density function $f_x(\cdot)$ of x_t, which is the Fourier transform of the ACF $\gamma_x(h)$ of x_t, that is
$$f_x(\nu) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma_x(h) \, \exp(-i\, h\, \pi\, \nu)$$
for $\nu \in [-1, \ 1]$, where $i = \sqrt{-1}$, satisfies
$$f_x(\nu) \sim \nu^{-2d} \quad \text{as} \quad \nu \to 0, \tag{4.2}$$
where $\nu \in [0, \ 1]$ denotes the frequency.

See the books by Hamilton (1994) and Brockwell and Davis (1991) for details about the spectral analysis. The basic idea and properties of the spectral density and its estimation are discussed in the section.

Of particular interest here is the behavior of ACF of x_t when $d < 0.5$. The property says that $\rho_x(h) \sim h^{2d-1}$, which decays at a polynomial, instead of exponential rate. For this reason, such an x_t process is called a long-memory time series. A special characteristic of the spectral density function in (4.2) is that the spectrum diverges to infinity as $\nu \to 0$. However, the spectral density function of a stationary ARMA process is bounded for all $\nu \in [-1,\ 1]$.

Earlier we used the binomial theorem for non-integer powers

$$(1-L)^d = \sum_{k=0}^{\infty} (-1)^k \binom{d}{k} L^k.$$

If the fractionally differenced series $(1-L)^d x_t$ follows an ARMA(p,q) model, then x_t is called an fractionally differenced autoregressive moving average (ARFIMA(p,d,q)) process, which is a generalized ARIMA model by allowing for non-integer d. In practice, if the sample ACF of a time series is not large in magnitude, but decays slowly, then the series may have long memory. For more discussions, we refer to the book by Beran (1994). For the pure fractionally differenced model in (4.1), one can estimate d using either a maximum likelihood method in the time domain (by assuming that the distribution is known) or the approximate Whittle likelihood (see below) or a regression method with logged periodogram at the lower frequencies (using (4.2)) in the frequency domain. Finally, long-memory models have attracted some attention in the finance literature in part because of the work on fractional Brownian motion in the continuous time models.

4.1.2 Spectral Density

We define the basic statistics used for detecting periodicities in time series and for determining whether different time series are related at certain frequencies. The power spectrum is a measure of the power or variance of a time series at a particular frequency ν; the function that displays the power or variance for each frequency ν, say $f_x(\nu)$ is called the power spectral density. Although the function $f_x(\cdot)$ is the Fourier transform of the autocovariance function $\gamma_x(h)$, we shall not make much use of this theoretical result in our generally applied discussion. For the theoretical results, please read the book by Brockwell and Davis (1991, §10.3).

The discrete Fourier transform of a sampled time series $\{x_t\}_{t=0}^{n-1}$ is defined as

$$X(\nu_k) = \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} x_t \, \exp\{-2\,\pi\,\nu_k\,t\}, \tag{4.3}$$

where $\nu_k = k/n$, $k = 0, \ldots, n-1$, defines the set of frequencies over which (4.3) is computed. This means that the frequencies over which (4.3) is evaluated are of the form $\nu = 0, 1/n, 2/n, \ldots, (n-1)/n$. The evaluation of (4.3) proceeds using the fast Fourier transform (DFT) and usually assumes that the length of the series, n is some power of 2. If the series is not a power of 2, it can be padded by adding zeros so that the extended series corresponds to the next highest power of two. One might want to do this anyway if frequencies of the form $\nu_k = k/n$ are not close enough to the frequencies of interest.

Since, the values of (4.3) have a real and imaginary part at every frequency and can be positive or negative, it is conventional to calculate first the squared magnitude of (4.3) which is called the periodogram. The periodogram is defined as

$$P_x(\nu_k) = |X(\nu_k)|^2 \tag{4.4}$$

and is just the sum of the squares of the sine and cosine transforms of the series. If we consider the power spectrum $f_x(\nu_k)$ as the quantity to estimate, it is approximately true[1]; see the books by Hamilton (1994) and Brockwell and Davis (1991, §10.3), that

$$E[P(\nu_k)] = E\left[|X(\nu_k)|^2\right] \approx f_x(\nu_k), \tag{4.5}$$

which implies that the periodogram is an approximately unbiased estimator of the power spectrum at frequency ν_k. One may also show that the approximate covariance between two frequencies ν_k and ν_l is zero for $k \neq l$ and both frequencies multiples of $1/n$. One can also show that the sine and cosine parts of (4.3) are uncorrelated and approximately normally distributed with equal variances $f_x(\nu_k)/2$. This implies that the squared standardized real and imaginary components have chi-squared distributions with 1 degree of freedom each. The periodogram is the sum of two such squared variables and must then have a chi-squared distribution with 2 degrees of freedom. In other words, the approximate distribution of $P(\nu_k)$ is exponential with the parameter $\lambda_k = f(\nu_k)$. The Whittle likelihood is based on the approximate exponential distribution of $\{P(\nu_k)\}_{k=0}^{n-1}$.

It should be noted that trends introduce apparent long range periodicities which inflate the spectrum at low frequencies (small values of ν_k) since the periodogram assumes that the trend is part of one very long frequency. This behavior will obscure possible important

[1] For Ph.D. students, I encourage you to read the related book to understand the details on this aspect, which is important for your future research.

components at higher frequencies. For this reason (as with the auto and cross correlations) it is important to work in (4.3) with the transform of the mean-adjusted $(x_t - \overline{x})$ or detrended $(x_t - \widehat{a} - \widehat{b}\,t)$ series. A further modification that sometimes improves the approximation in the expectation in (4.5) is tapering, a procedure where by each point x_t is replaced by $a_t\, x_t$, where a_t is a function, often a cosine bell, that is maximum at the midpoint of the series and decays away at the extremes. Tapering makes a difference in engineering applications where there are large variations in the value of the spectral density $f_x(\cdot)$. As a final comment, we note that the fast Fourier transform algorithms for computing (4.3) generally work best when the sample length is some power of two. For this reason, it is conventional to extend the data series x_t, $t = 0, 1, \ldots, n-1$ to a length $n^* > n$ that is a power of two by replacing the missing values by $x_t = 0$, $t = n+1, \ldots, n^*-1$. If this is done, the frequency ordinates $\nu_k = k/n$ are replaced by $\nu_k^* = k/n^*$ and we proceed as before.

Theorem 10.3.2 in Brockwell and Davis (1991, p.347) shows that $P(\nu_k)$ is not a consistent estimator of $f_x(\nu)$. Since for large n, the periodogram ordinates are approximately uncorrelated with variances changing only slightly over small frequency intervals, we might hope to construct a consistent estimator of $f_x(\nu)$ by averaging the periodogram ordinates in a small neighborhood of ν, which is called the smoothing estimator. On the other hand, such a smoothing estimator can reduce the variability of the periodogram estimator of the previous section. This procedure leads to a spectral estimator of the form

$$\widehat{f}_x(\nu_k) = \frac{1}{2L+1} \sum_{l=-L}^{L} P_x(\nu_k + l/n) = \sum_{l=-L}^{L} w_l\, P_x(\nu_k + l/n), \tag{4.6}$$

when the periodogram is smoothed over $2L+1$ frequencies. The width of the interval over which the frequency average is taken is called the bandwidth. Since there are $2L+1$ frequencies of width $1/n$, the bandwidth B in this case is approximately $B = (2L+1)/n$. The smoothing procedure improves the quality of the estimator for the spectrum since it is now the average of L random variables each having a chi-squared distribution with 2 degrees of freedom. The distribution of the smoothed estimator then will have $df = 2(2L+1)$ degrees of freedom. If the series is adjusted to the next highest power of 2, say n^*, the adjusted degrees of freedom for the estimator will be $df^* = 2(2L+1)n/n^*$ and the new bandwidth will be $B = (2L+1)/n^*$.

Clearly, in (4.6), $w_l = 1/(2L+1)$, which is called the Daniell window (the corresponding spectral window is given by the Daniell kernel). One popular approach is to take w_l to be

$$w_l = w(\nu_l)/n.$$

If $w(x) = r_n^{-1}\sin^2(r\,x/2)/\sin^2(x/2)$, it is the well known Barlett or triangular window (the corresponding spectral window is given by the Fejer kernel), where $r_n = \lfloor n/(2L)\rfloor$. If $w(x) = \sin((r+0.5)\,x)/\sin(x/2)$, it is the rectangular window (the corresponding spectral window is given by the Dirichlet kernel). See Brockwell and Davis (1991, §10.4) for more discussions.

4.1.3 Applications

The usage of the function **fracdiff()** is

```
fracdiff(x, nar = 0, nma = 0,
         ar = rep(NA, max(nar, 1)), ma = rep(NA, max(nma, 1)),
         dtol = NULL, drange = c(0, 0.5), h, M = 100)
```

This function can be used to compute the maximum likelihood estimators of the parameters of a fractionally-differenced ARIMA(p, d, q) model, together (if possible) with their estimated covariance and correlation matrices and standard errors, as well as the value of the maximized likelihood. The likelihood is approximated using the fast and accurate method of Haslett and Raftery (1989). To generate simulated long-memory time series data from the fractional ARIMA(p, d, q) model, we can use the following function **fracdiff.sim()** and its usage is

```
fracdiff.sim(n, ar = NULL, ma = NULL, d,
             rand.gen = rnorm, innov = rand.gen(n+q, ...),
            n.start = NA, allow.0.nstart = FALSE, ..., mu = 0.)
```

An alternative way to simulate a long memory time series is to use the function **arima.sim()**. The menu for the package **fracdiff** can be downloaded from the web site at
http://cran.cnr.berkeley.edu/doc/packages/fracdiff.pdf

The function **spec.pgram()** in **R** calculates the periodogram using a fast Fourier transform, and optionally smooths the result with a series of modified Daniell smoothers (moving averages giving half weight to the end values). The usage of this function is

```
spec.pgram(x, spans = NULL, kernel, taper = 0.1,
           pad = 0, fast = TRUE, demean = FALSE, detrend = TRUE,
```

```
                plot = TRUE, na.action = na.fail, ...)
```

We can also use the function **spectrum()** to estimate the spectral density of a time series and its usage is

```
spectrum(x, ..., method = c("pgram", "ar"))
```

Finally, it is worth to pointing out that there is a package called **longmemo** for long-memory processes, which can be downloaded from
http://cran.cnr.berkeley.edu/doc/packages/longmemo.pdf. This package also provides a simple periodogram estimation by function **per()** and other functions like **llplot()** and **lxplot()** for making graphs for spectral density. See the menu for details.

Example 9.1: As an illustration, Figure 4.1 show the sample ACFs of the absolute series of daily simple returns for the CRSP value-weighted (left top panel) and equal-weighted (right top panel) indexes from July 3, 1962 to December 31, 1997 and the sample partial autocorrelation function of the absolute series of daily simple returns for the CRSP value-weighted (left middle panel) and equal-weighted (right middle panel) indexes. The ACFs are relatively small in magnitude, but decay very slowly; they appear to be significant at the 5% level even after 300 lags. There are only the first few lags for PACFs outside the confidence interval and then the rest is basically within the confidence interval. For more information about the behavior of sample ACF of absolute return series, see Ding, Granger, and Engle (1993). To estimate the long memory parameter estimate d, we can use the function **fracdiff()** in the package **fracdiff** in **R** and results are $\widehat{d} = 0.1867$ for the absolute returns of the value-weighted index and $\widehat{d} = 0.2732$ for the absolute returns of the equal-weighted index. To support our conclusion above, we plot the log smoothed spectral density estimation of the absolute series of daily simple returns for the CRSP value-weighted (left bottom panel) and equal-weighted (right bottom panel). They show clearly that both log spectral densities decay like a log function and they support the spectral densities behavior like (4.2).

4.2 Related Problems and New Developments

The reading materials are the papers by Ding, Granger, and Engle (1993), Krämer, Sibbertsen and Kleiber (2002), Zeileis, Leisch, Hornik, and Kleiber (2002), and Sibbertsen (2004a).

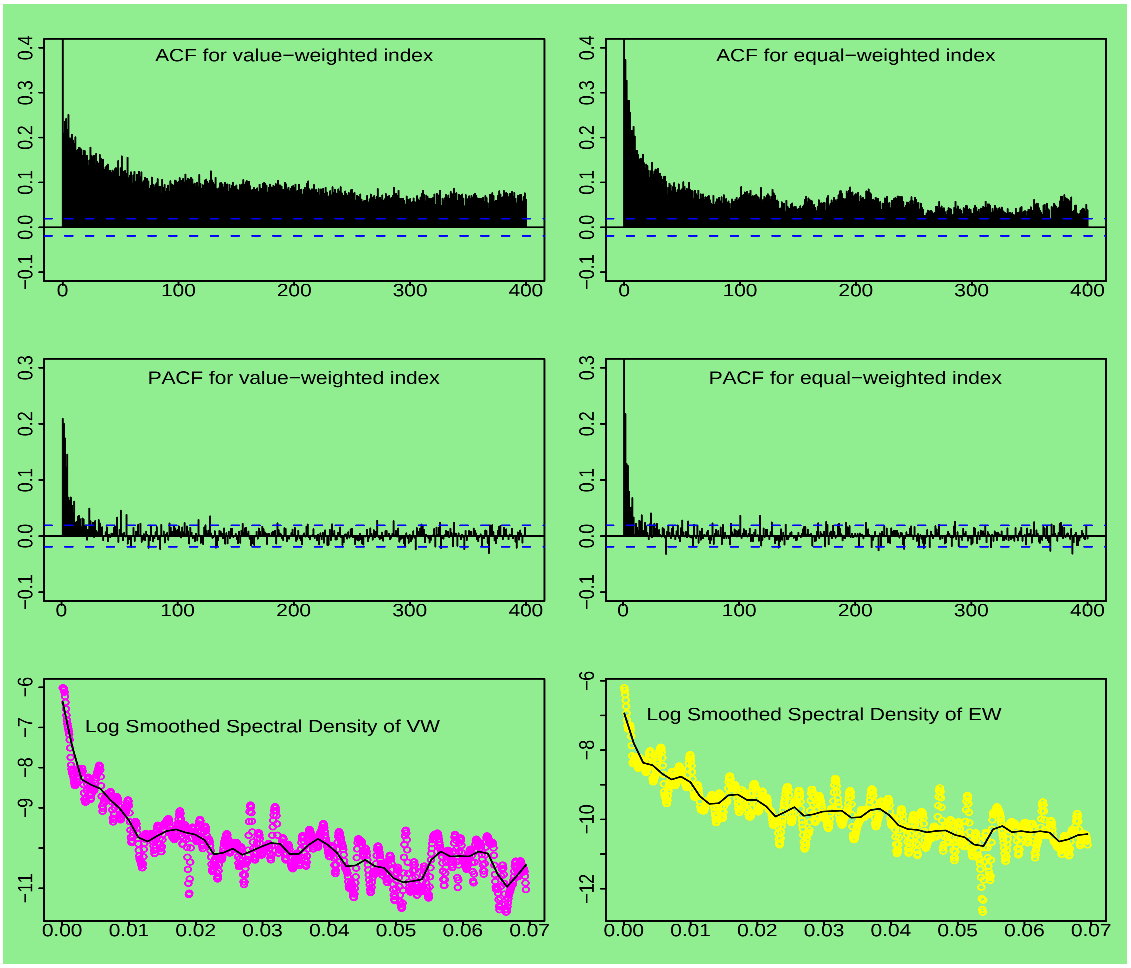

Figure 4.1: Sample autocorrelation function of the absolute series of daily simple returns for the CRSP value-weighted (left top panel) and equal-weighted (right top panel) indexes. Sample partial autocorrelation function of the absolute series of daily simple returns for the CRSP value-weighted (left middle panel) and equal-weighted (right middle panel) indexes. The log smoothed spectral density estimation of the absolute series of daily simple returns for the CRSP value-weighted (left bottom panel) and equal-weighted (right bottom panel) indexes.

4.2.1 Long Memory versus Structural Breaks

It is a well known stylized fact that many financial time series such as squares or absolute values of returns or volatilities, even returns themselves behave as if they had long memory; see Ding, Granger and Engle (1993) and Sibbertsen (2004b). On the other hand, it is also well known that long memory is easily confused with **structural change**, in the sense that the slow decay of empirical autocorrelations which is typical for a time series with long memory is also produced when a shortmemory time series exhibits **structural breaks**. Therefore it is of considerable theoretical and empirical interest to discriminate between these sources of slowly decaying empirical autocorrelations.

Structural break is another type of nonstationarity arises when the population regression function changes over the sample period. This may occur because of changes in economic policy, changes in the structure of the economy or industry, events that change the dynamics of specific industries or firm related quantities such as inventories, sales, and production, etc. If such changes, called breaks, occur then regression models that neglect those changes lead to a misleading inference or forecasting.

Breaks may result from a discrete change (or changes) in the population regression coefficients at distinct dates or from a gradual evolution of the coefficients over a longer period of time. Discrete breaks may be a result of some major changes in economic policy or in the economy (oil shocks) while "gradual" breaks, population parameters evolve slowly over time, may be a result of slow evolution of economic policy. The former might be characterized by an indicator function and latter can be described by a smooth transition function.

If a break occurs in the population parameters during the sample, then the OLS regression estimates over the full sample will estimate a relationship that holds on **"average"**. Now the question is how to test breaks.

4.2.2 Testing for Breaks (Instability)

Tests for breaks in the regression parameters depend on whether the break date is know or not. If the date of the hypothesized break in the coefficients is known, then the null hypothesis of no break can be testing using a dummy or indicator variable. For example, consider the following model:

$$y_t = \begin{cases} \beta_0 + \beta_1 y_{t-1} + \delta_1 \, x_t + u_t, & \text{if } t \le \tau, \\ (\beta_0 + \gamma_0) + (\beta_1 + \gamma_1) y_{t-1} + (\delta_1 + \gamma_2) \, x_t + u_t, & \text{if } t > \tau, \end{cases}$$

where τ denotes the hypothesized break date. Under the null hypothesis of no break, H_0 : $\gamma_0 = \gamma_1 = \gamma_2 = 0$, and the hypothesis of a break can be tested using the F-statistic. This is called a **Chow test** proposed by Chow (1960) for a break at a known break date. Indeed, the above structural break model can be regarded as a special case of the following trending time series model

$$y_t = \beta_0(t) + \beta_1(t) \, y_{t-1} + \delta_1(t) \, x_t + u_t.$$

For more discussions, see Cai (2007). If there is a distinct break in the regression function, the date at which the largest Chow statistic occurs is an estimator of the break date.

If there are more variables or more lags, this test can be extended by constructing binary variable interaction variables for all the dependent variables. This approach can be modified to check for a break in a subset of the coefficients. The break date is unknown in most of the applications but you may suspect that a break occurred sometime between two dates, τ_0 and τ_1. The Chow test can be modified to handle this by testing for break at all possible dates t in between τ_0 and τ_1, then using the largest of the resulting F-statistics to test for a break at an unknown date. This modified test is often called *Quandt likelihood ratio* (QLR) statistic or the *sup Wald* or **sup***F* statistic:

$$\mathrm{sup}\mathbf{F} = \max\{F(\tau_0),\ F(\tau_0+1), \cdots, F(\tau_1)\}.$$

Since the sup$F$ statistic is the largest of many F-statistics, its distribution is not the same as an individual F-statistic. The critical values for sup$F$ statistic must be obtained from a special distribution. This distribution depends on the number of restriction being tested, m, τ_0, τ_1, and the subsample over which the F-statistics are computed expressed as a fraction of the total sample size. Other types of F-tests are the average F and exponential F given by

$$\mathrm{ave}\mathbf{F} = \frac{1}{\tau_1 - \tau_0 + 1} \sum_{j=\tau_0}^{\tau_1} F_j$$

and

$$\mathrm{exp}\mathbf{F} = \log\left(\frac{1}{\tau_1 - \tau_0 + 1} \sum_{j=\tau_0}^{\tau_1} \exp(F_j/2)\right).$$

For details on modified F-tests, see the papers by Hansen (1992) and Andrews (1993).

For the large-sample approximation to the distribution of the sup$F$ statistic to be a good one, the subsample endpoints, τ_0 and τ_1, can not be too close to the end of the sample. That is why the sup$F$ statistic is computed over a "trimmed" subset of the sample. A popular choice is to use 15% trimming, that is, to set for $\tau_0 = 0.15T$ and $\tau_1 = 0.85T$. With 15% trimming, the F-statistic is computed for break dates in the central 70% of the sample. Table 4.1 presents the critical values for sup$F$ statistic computed with 15% trimming. This table is from Stock and Watson (2003) and you should check the book for a complete table. The

Table 4.1: Critical Values of the QLR statistic with 15% Trimming

Number of restrictions (m)	10%	5%	1%
1	7.12	8.68	12.16
2	5.00	5.86	7.78
3	4.09	4.71	6.02
4	3.59	4.09	5.12
5	3.26	3.66	4.53
6	3.02	3.37	4.12
7	2.84	3.15	3.82
8	2.69	2.98	3.57
9	2.58	2.84	3.38
10	2.48	2.71	3.23

sup$F$ test can detect a single break, multiple discrete breaks, and a slow evolution of the regression parameters.

Three classes of structural change tests (or tests for parameter instability) which have been receiving much attention in both the statistics and econometrics communities but have been developed in rather loosely connected lines of research are unified by embedding them into the framework of generalized M-fluctuation tests (Zeileis and Hornik (2003)). These classes are tests based on maximum likelihood scores (including the Nyblom-Hansen test), on F statistics (supF, aveF, expF tests) and on OLS residuals (**OLS-based CUSUM** and **MOSUM tests**; see Chu, Hornik and Kuan (1995), which is a special case of the so called empirical fluctuation process, termed as **efp** in Zeileis, Leisch, Hornik, and Kleiber (2002) and Zeileis and Hornik (2003)). Zeileis (2005) showed that representatives from these classes are special cases of the generalized M-fluctuation tests, based on the same functional central limit theorem, but employing different functionals for capturing excessive fluctuations. After embedding these tests into the same framework and thus understanding the relationship between these procedures for testing in historical samples, it is shown how the tests can also be extended to a monitoring situation. This is achieved by establishing a general M-fluctuation monitoring procedure and then applying the different functionals corresponding to monitoring with ML scores, F statistics and OLS residuals. In particular, an extension of the supF test to a monitoring scenario is suggested and illustrated on a real world data set.

In **R**, there are two packages **strucchange** developed by Zeileis, Leisch, Hornik, and

Kleiber (2002) and **segmented** to provide several testing methods for testing breaks. They can be downloaded at **http://cran.cnr.berkeley.edu/doc/packages/strucchange.pdf** and **http://cran.cnr.berkeley.edu/doc/packages/segmented.pdf**, respectively.

Example 9.2: We use the data build-in in **R** for the minimal water flow of the Nile River when planning the Ashwan Dam; see Hurst (1951), yearly data from 1871 to 1970 with 100 observations and the data name **Nile**. Also, we might use the data build-in the package **longmemo** in **R**, yearly data with 633 observations from 667 to 1284 and the data name **NileMin**. Here we use **R** to run the data set in **Nile**. As a result, the p-value for testing break

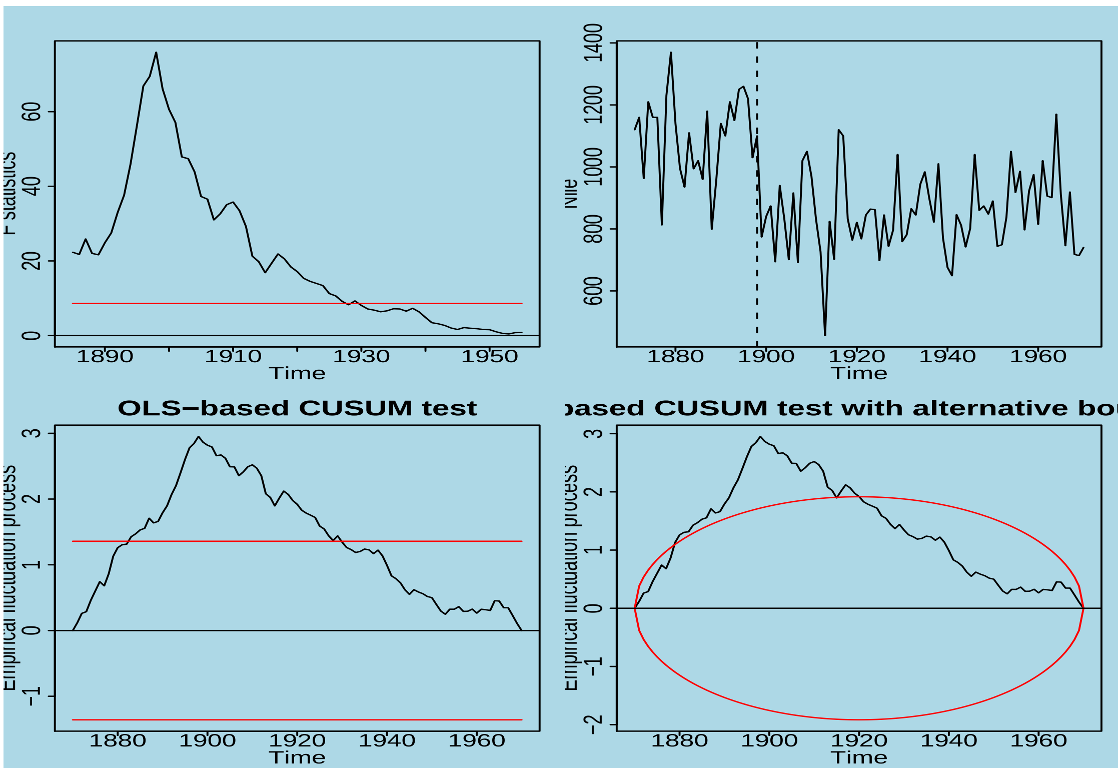

Figure 4.2: Break testing results for the Nile River data: (a) Plot of F-statistics. (b) The scatterplot with the breakpoint. (c) Plot of the empirical fluctuation process with linear boundaries. (d) Plot of the empirical fluctuation process with alternative boundaries.

is very small (see the computer output) so that H_0 is rejected. The details are summarized in Figures 4.2(a)-(b). It is clear from Figure 4.2(b) that there is one breakpoint for the Nile River data: the annual flows drop in 1898 because the first Aswan dam was built. To test the null hypothesis that the annual flow remains constant over the years, we also compute OLS-based CUSUM process and plot with standard linear and alternative boundaries in Figures 4.2(c)-(d).

Example 9.3: We build a time series model for real oil price (quarterly) listed in the tenth column in file *"ex9-3.csv"*, ranged from the first quarter of 1959 to the third quarter of 2002. Before we build such a time series model, we want to see if there is any structure change for oil price. As a result, the p-value for testing break is very small (see the computer output) so that H_0 is rejected. The details are summarized in Figure 4.3. It is clear from Figure

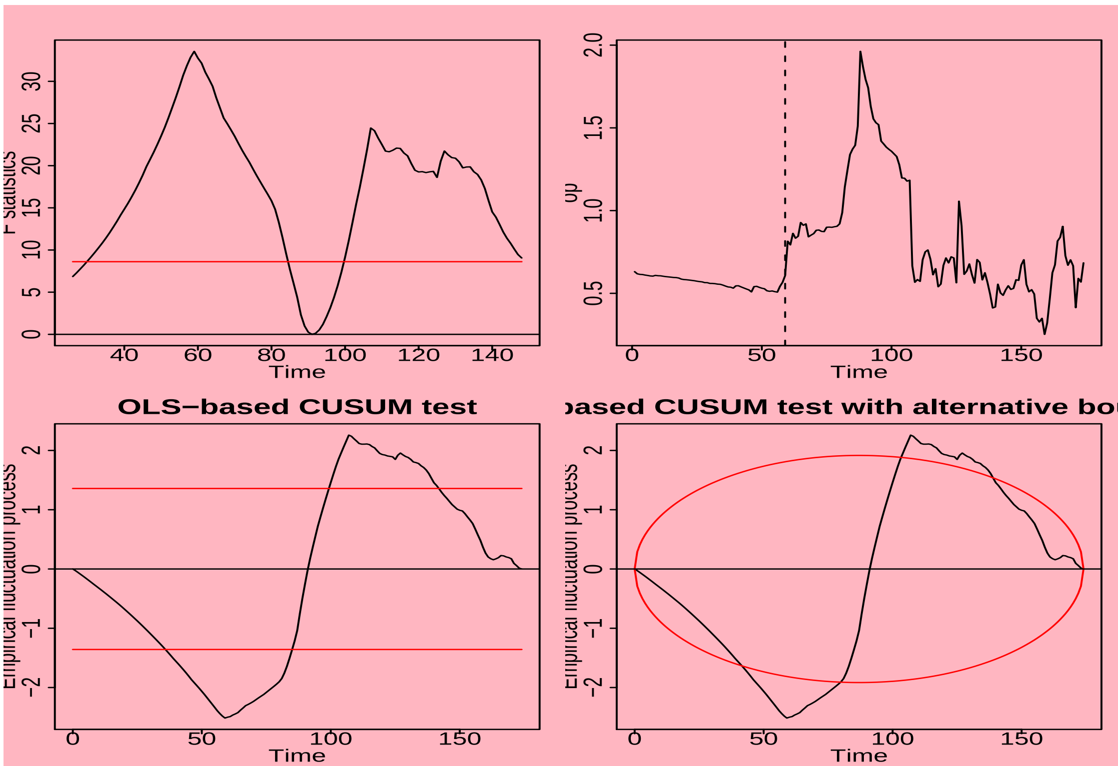

Figure 4.3: Break testing results for the oil price data: (a) Plot of F-statistics. (b) Scatterplot with the breakpoint. (c) Plot of the empirical fluctuation process with linear boundaries. (d) Plot of the empirical fluctuation process with alternative boundaries.

4.3(b) that there is one breakpoint for the oil price. We also compute OLS-based CUSUM process and plot with standard linear and alternative boundaries in Figures 4.3(c)-(d).

If we consider the quarterly price level (CPI) listed in the eighth column in file *"ex9-3.csv"* (1959.1 – 2002.3), we want to see if there is any structure change for quarterly consumer price index. The details are summarized in Figure 4.4. It is clear from Figure 4.4(b) that there would be one breakpoint for the consumer price index. We also compute OLS-based CUSUM process and plot with standard linear and alternative boundaries in Figures 4.4(c)-(d). **Please thank about the conclusion for CPI!!! Do you believe this result? If not, what happen?**

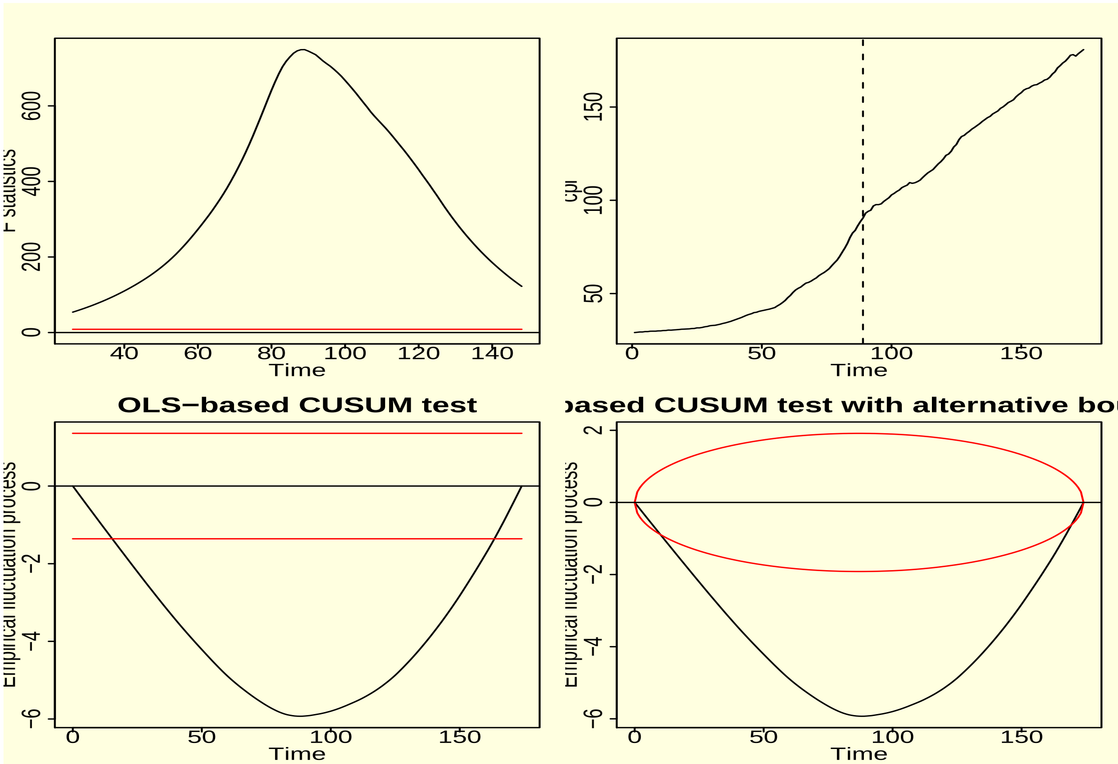

Figure 4.4: Break testing results for the consumer price index data: (a) Plot of F-statistics. (b) Scatterplot with the breakpoint. (c) Plot of the empirical fluctuation process with linear boundaries. (d) Plot of the empirical fluctuation process with alternative boundaries.

Sometimes, you would suspect that a series may either have a unit root or be a trend stationary process that has a structural break at some unknown period of time and you would want to test the null hypothesis of unit root against the alternative of a trend stationary process with a structural break. This is exactly the hypothesis testing procedure proposed by Zivot and Andrews (1992). In this testing procedure, the null hypothesis is a unit root process without any structural breaks and the alternative hypothesis is a trend stationary process with possible structural change occurring at an unknown point in time. Zivot and Andrews (1992) suggested estimating the following regression:

$$x_t = \begin{cases} \mu + \beta\, t + \alpha\, x_{t-1} + \sum_{i=1}^{k} c_i\, \Delta\, x_{t-i} + e_t, & \text{if } t \le \tau\, T, \\ [\mu + \theta] + [\beta\, t + \gamma(t - T^B)] + \alpha\, x_{t-1} + \sum_{i=1}^{k} c_i\, \Delta\, x_{t-i} + e_t, & \text{if } t > \tau\, T, \end{cases} \tag{4.7}$$

where $\tau = T^B/T$ is the break fraction. Model (4.7) is estimated by OLS with the break points ranging over the sample and the t-statistic for testing $\alpha = 1$ is computed. The minimum t-statistic is reported. Critical values for 1%, 5% and 10% critical values are -5.34, -4.8 and -4.58, respectively. The appropriate number of lags in differences is estimated for each value of τ. Please read the papers by Zivot and Andrews (1992) and Sadorsky (1999) for more details about this method and empirical applications.

4.2.3 Long Memory versus Trends

A more general problem than distinguishing long memory and structural breaks is in a way the question if general trends in the data can cause the Hurst effect. The paper by Bhattacharya et al (1983) was the first to deal this problem. They found that adding a **deterministic trend** to a short memory process can cause **spurious** long memory. Now the problem becomes more complicated for modeling long memory time series. For this issue, we will follow Section 5 of Zeileis (2004a). Note that there are a lot of ongoing research (theoretical and empirical) works in this area.

4.3 Computer Codes

```
#####################
# This is Example 9.1
#####################
z1<-matrix(scan("c:/res-teach/xiamen12-06/data/ex9-1.txt"),
byrow=T,ncol=5)
vw=abs(z1[,3])
n_vw=length(vw)
ew=abs(z1[,4])
postscript(file="c:/res-teach/xiamen12-06/figs/fig-9.1.eps",
 horizontal=F,width=6,height=6)
par(mfrow=c(3,2),mex=0.4,bg="light green")
acf(vw, ylab="",xlab="",ylim=c(-0.1,0.4),lag=400,main="")
text(200,0.38,"ACF for value-weighted index")
acf(ew, ylab="",xlab="",ylim=c(-0.1,0.4),lag=400,main="")
text(200,0.38,"ACF for equal-weighted index")
pacf(vw, ylab="",xlab="",ylim=c(-0.1,0.3),lag=400,main="")
text(200,0.28,"PACF for value-weighted index")
pacf(ew, ylab="",xlab="",ylim=c(-0.1,0.3),lag=400,main="")
text(200,0.28,"PACF for equal-weighted index")
```

```
library(fracdiff)
d1=fracdiff(vw,ar=0,ma=0)
d2=fracdiff(ew,ar=0,ma=0)
print(c(d1$d,d2$d))
m1=round(log(n_vw)/log(2)+0.5)
pad1=1-n_vw/2^m1
vw_spec=spec.pgram(vw,spans=c(5,7),demean=T,detrend=T,pad=pad1,plot=F)
ew_spec=spec.pgram(ew,spans=c(5,7),demean=T,detrend=T,pad=pad1,plot=F)
vw_x=vw_spec$freq[1:1000]
vw_y=vw_spec$spec[1:1000]
ew_x=ew_spec$freq[1:1000]
ew_y=ew_spec$spec[1:1000]
scatter.smooth(vw_x,log(vw_y),span=1/15,ylab="",xlab="",col=6,cex=0.7,
  main="")
text(0.03,-7,"Log Smoothed Spectral Density of VW")
scatter.smooth(ew_x,log(ew_y),span=1/15,ylab="",xlab="",col=7,cex=0.7,
  main="")
text(0.03,-7,"Log Smoothed Spectral Density of EW")
dev.off()

##########################
# This is for Example 9.2
##########################
graphics.off()
library(strucchange)
library(longmemo)         # not used

postscript(file="c:/res-teach/xiamen12-06/figs/fig-9.2.eps",
 horizontal=F,width=6,height=6)
par(mfrow=c(2,2),mex=0.4,bg="light blue")
#if(! "package:stats" %in% search()) library(ts)
## Nile data with one breakpoint: the annual flows drop in 1898
```

```
## because the first Ashwan dam was built
data(Nile)
## test whether the annual flow remains constant over the years
fs.nile=Fstats(Nile ~ 1)
plot(fs.nile)
print(sctest(fs.nile))
plot(Nile)
lines(breakpoints(fs.nile))

## test the null hypothesis that the annual flow remains constant
## over the years
## compute OLS-based CUSUM process and plot
## with standard and alternative boundaries
ocus.nile=efp(Nile ~ 1, type = "OLS-CUSUM")
plot(ocus.nile)
plot(ocus.nile, alpha = 0.01, alt.boundary = TRUE)
## calculate corresponding test statistic
print(sctest(ocus.nile))
dev.off()

#########################
# This is for Example 9.3
#########################

y=read.csv("c:/res-teach/xiamen12-06/data/ex9-3.csv",header=T,skip=1)
op=y[,10]                         # oil price
postscript(file="c:/res-teach/xiamen12-06/figs/fig-9.3.eps",
 horizontal=F,width=6,height=6)
par(mfrow=c(2,2),mex=0.4,bg="light pink")
op=ts(op)
fs.op=Fstats(op ~ 1)               # no lags and covariate
#plot(op,type="l")
```

```
plot(fs.op)
print(sctest(fs.op))
## visualize the breakpoint implied by the argmax of the F statistics
plot(op,type="l")
lines(breakpoints(fs.op))
ocus.op=efp(op~ 1, type = "OLS-CUSUM")
plot(ocus.op)
plot(ocus.op, alpha = 0.01, alt.boundary = TRUE)
## calculate corresponding test statistic
print(sctest(ocus.op))
dev.off()
cpi=y[,8]
cpi=ts(cpi)
fs.cpi=Fstats(cpi~1)
print(sctest(fs.cpi))
postscript(file="c:/res-teach/xiamen12-06/figs/fig-9.4.eps",
 horizontal=F,width=6,height=6)
#win.graph()
par(mfrow=c(2,2),mex=0.4,bg="light yellow")
#plot(cpi,type="l")
plot(fs.cpi)
## visualize the breakpoint implied by the argmax of the F statistics
plot(cpi,type="l")
lines(breakpoints(fs.cpi))
ocus.cpi=efp(cpi~ 1, type = "OLS-CUSUM")
plot(ocus.cpi)
plot(ocus.cpi, alpha = 0.01, alt.boundary = TRUE)
## calculate corresponding test statistic
print(sctest(ocus.cpi))
dev.off()
```

4.4 References

Andrews, D.W.K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, **61**, 821-856.

Bhattacharya, R.N., V.K. Gupta and W. Waymire (1983). The Hurst effect under trends. *Journal of Applied Probability*, **20**, 649-662.

Beran, J. (1994). *Statistics for Long-Memory Processes*. Chapman and Hall, London.

Brockwell, P.J. and Davis, R.A. (1991). *Time Series Theory and Methods*. New York: Springer.

Chan, N.H. and C.Z. Wei (1988). Limiting distributions of least squares estimates of unstable autoregressive processes. *Annals of Statistics*, **16**, 367-401.

Chow, G.C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, **28**, 591-605.

Chu, C.S.J, K. Hornik and C.M. Kuan (1995). MOSUM tests for parameter constancy. *Biometrika*, **82**, 603-617.

Ding, Z., C.W.J. Granger and R.F. Engle (1993). A long memory property of stock returns and a new model. *Journal of Empirical Finance*, **1**, 83-106.

Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press.

Hansen, B.E. (1992). Testing for parameter instability in linear models. *Journal of Policy Modeling*, **14**, 517-533.

Haslett, J. and A.E. Raftery (1989). Space-time modelling with long-memory dependence: Assessing Ireland's wind power resource (with discussion). *Applied Statistics*, **38**, 1-50.

Hsu, C.-C., and C.-M. Kuan. (2000). Long memory or structural changes: Testing method and empirical examination. *Working paper*, Department of Economics, National Central University, Taiwan.

Hurst, H.E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, **116**, 770-799.

Granger, C.W.J. (1966). The typical spectral shape of an economic variable. *Econometrica*, **34**, 150-161.

Krämer, W., P. Sibbertsen and C. Kleiber (2002). Long memory versus structural change in financial tim series. *Allgemeines Statistisches Archiv*, **86**, 83-96.

Nyblom, J. (1989). Testing for the constancy of parameters over time. *Journal of the American Statistical Association*, **84**, 223-230.

Sadorsky, P. (1999). Oil price shocks and stock market activity. *Energy Economics*, **21**, 449-469.

Sibbertsen, P. (2004a). Long memory versus structural breaks: An overview. *Statistical Papers*, **45**, 465-515.

Sibbertsen, P. (2004b). Long memory in volatilities of German stock returns. *Empirical Economics*, **29**, 477-488.

Stock, J.H. and M.W. Watson (2003). *Introduction to Econometrics*. Addison-Wesley.

Tiao, G.C. and R.S. Tsay (1983). Consistency properties of least squares estimates of autoregressive parameters in ARMA models. *Annals of Statistics*, **11**, 856-871.

Tsay, R.S. (2005). *Analysis of Financial Time Series*, 2th Edition. John Wiley & Sons, New York.

Zeileis, A. (2005). A unified approach to structural change tests based on ML scores, F statistics, and OLS residuals. *Econometric Reviews*, **24**, 445-466.

Zeileis A. and Hornik, K. (2003). Generalized M-fluctuation tests for parameter instability. *Report 80*, SFB "Adaptive Information Systems and Modelling in Economics and Management Science". URL **http://www.wu-wien.ac.at/am/reports.htm#80**.

Zeileis, A., F. Leisch, K. Hornik and C. Kleiber (2002). *strucchange:* An R package for testing for structural change in linear regression models. *Journal of Statistical Software*, **7**, Issue 2, 1-38.

Zivot, E. and D.W.K. Andrews (1992). Further evidence on the great crash, the oil price shock and the unit root hypothesis. *Journal of Business and Economics Statistics*, **10**, 251-270.