INSTRUCTOR'S RESOURCE GUIDE TO ACCOMPANY

APPLIED ECONOMETRIC TIME SERIES (2nd edition)

Walter Enders
University of Alabama

Prepared by

Pin Chung
American Reinsurance Company and Iowa State University

Walter Enders
University of Alabama

Ling Shao
University of Alabama

Jingan Yuan
University of Alabama

PREFACE

This Instructor's Manual is designed to accompany the second edition of Walter Enders' *Applied Econometric Time Series* (*AETS*). As in the first edition, the text instructs by induction. The method is to take a simple example and build towards more general models and econometric procedures. A large number of examples are included in the body of each chapter. Many of the mathematical proofs are performed in the text and detailed examples of each estimation procedure are provided. The approach is one of learning-by-doing. As such, the mathematical questions and the suggested estimations at the end of each chapter are important. In addition, it is useful to have students perform the type of semester project described at the end of this manual.

One aim of this manual is to provide the answers to each of the mathematical problems. Many of these questions are answered in great detail. Our goal was *not* to provide the most mathematically elegant solution techniques. Sometimes a long and drawn-out answer provides more insight than a concise proof.

This second aim is to provide sample programs that can be used to obtain the results reported in the 'Questions and Exercises' sections of *AETS*. Students should be encouraged to work through as many of these exercises as possible. In order to work through the exercises, it is necessary to have access to a statistical package such as EViews, Microfit, PC-GIVE, or RATS, SAS, SHAZAM or STATA. Matrix packages such as MATLAB, and GAUSS are not as convenient for univariate models. Some of these packages, such as EViews, allow you to perform most of the exercises using pull-down menus. Others, such as GAUSS, need to be programmed to perform relatively simple tasks. It is not possible to include programs for each of these packages within this small manual. There were several factors leading me to provide programs written for RATS and STATA. First, the RATS Programming Manual can be downloaded (at no charge) from www.estima.com/enders. The Programming Manual provides a complete discussion of many of the programming tasks used in time-series econometrics. STATA was included since it is a popular package that most would not consider to be a time-series package. Nevertheless, as shown below, STATA can produce almost all of the results obtained in the text. Adobe Acrobat allows you to copy a program from the *.pdf version of this manual and paste it directly into STATA or RATS. The languages used in RATS and STATA are relatively transparent. As such, users of other packages should be able to translate the programs provided here.

As stated in the Preface of *AETS*, the text is certain to contain a number of errors. If the first edition is any guide, the number is embarrassingly large. I will keep a list of typos and corrections on my Web page: www.cba.ua.edu/~wenders. Moreover, time-series methods and techniques keep evolving very rapidly. I will try to keep you updated by posting research notes and clarifications on my Web page. I would be happy to post any useful programs or communications you might have; my e-mail address is wenders@cba.ua.edu.

CONTENTS

CHAPTER 1

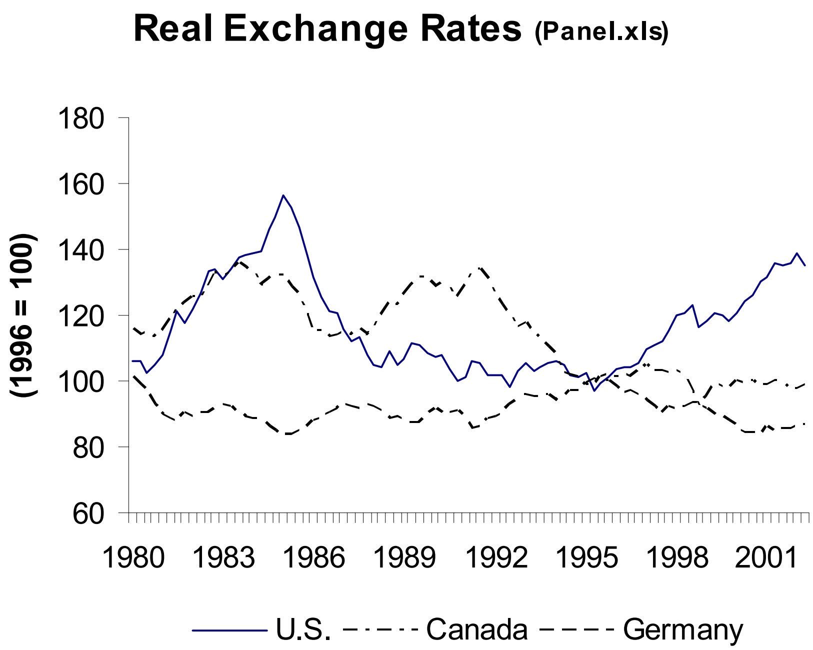
DIFFERENCE EQUATIONS

Lecture Suggestions

Nearly all students will have some familiarity the concepts developed in the chapter. A first course in integral calculus makes reference to convergent versus divergent solutions. I draw the analogy between the particular solution to a difference equation and indefinite integrals.

It is important to stress the distinction between convergent and divergent solutions. Be sure to emphasize the relationship between characteristic roots and the convergence or divergence of a sequence. Much of the current time-series literature focuses on the issue of unit roots. It is wise to introduce students to the properties of difference equations with unitary characteristic roots at this early stage in the course. Question 5 at the end of this chapter is designed to preview this important issue. The tools to emphasize are the **method of undetermined coefficients** and **lag operators**. Few students will have been exposed to these methods in other classes.

I use overheads to show the students several data series and ask them to discuss the type of difference equation model that might capture the properties of each. The figure below shows three of the real exchange rate series used in Chapter 4. Some students see a tendency for the series to revert to a long-run mean value. The classroom discussion might center on the appropriate way to model the tendency for the levels to meander. At this stage, the precise models are not important. The objective is for students to conceptualize economic data in terms of difference equations.



Answers to Questions

1. Consider the difference equation: $y_t = a_0 + a_1 y_{t-1}$ with the initial condition y_0. Jill solved the difference equation by iterating backwards:

$$\begin{aligned} y_t &= a_0 + a_1 y_{t-1} \\ &= a_0 + a_1[a_0 + a_1 y_{t-2}] \\ &= a_0 + a_0 a_1 + a_0(a_1)^2 + ... + a_0(a_1)^{t-1} + (a_1)^t y_0 \end{aligned}$$

Bill added the homogeneous and particular solutions to obtain: $y_t = a_0/(1 - a_1) + (a_1)^t[y_0 - a_0/(1 - a_1)]$.

A. Show that the two solutions are identical for $|a_1| < 1$.

Answer: The key is to demonstrate:

$$a_0 + a_0 a_1 + a_0(a_1)^2 + ... + a_0(a_1)^{t-1} + (a_1)^t y_0 = a_0/(1 - a_1) + (a_1)^t[y_0 - a_0/(1 - a_1)]$$

First, cancel $(a_1)^t y_0$ from each side and then divide by a_0. The two sides of the equation are identical if:

$$1 + a_1 + (a_1)^2 + ... + (a_1)^{t-1} = 1/(1 - a_1) - (a_1)^t/(1 - a_1)$$

Now, multiply each side by $(1 - a_1)$ to obtain:

$$(1 - a_1)[1 + a_1 + (a_1)^2 + ... + (a_1)^{t-1}] = 1 - (a_1)^t$$

Multiply the two expressions in parentheses to obtain:

$$1 - (a_1)^t = 1 - (a_1)^t$$

The two sides of the equation are identical. Hence, Jill and Bob obtained the identical answer.

B. Show that for $a_1 = 1$, Jill's solution is equivalent to: $y_t = a_0t + y_0$. How would you use Bill's method to arrive at this same conclusion in the case $a_1 = 1$.

Answer: When $a_1 = 1$, Jill's solution can be written as:

$$\begin{aligned} y_t &= a_0(1^0 + 1^1 + 1^2 + ... + 1^{t-1}) + y_0 \\ &= a_0t + y_0 \end{aligned}$$

To use Bill's method, find the homogeneous solution from the equation $y_t = y_{t-1}$. Clearly, the homogeneous solution is any arbitrary constant A. The key in finding the particular solution is to realize that the characteristic root is unity. In this instance, the particular solution has the form a_0t. Adding the homogeneous and particular solutions, the general solution is

$$y_t = a_0t + A$$

To eliminate the arbitrary constant, impose the initial condition. The general solution must hold for all t including $t = 0$. Hence, at $t = 0$, $y_0 = a_0t + A$ so that $A = y_0$. Hence, Bill's method yields:

$$\boxed{y_t = a_0t + y_0}$$

2. The Cobweb model in section 5 assumed *static* price expectations. Consider an alternative formulation called *adaptive expectations*. Let the expected price in t (denoted by p_t^*) be a weighted average of the price in t-1 and the price expectation of the previous period. Formally:

$$p_t^* = \alpha p_{t-1} + (1 - \alpha)\, p_{t-1}^* \qquad\qquad 0 < \alpha \leq 1.$$

Clearly, when $\alpha = 1$, the static and adaptive expectations schemes are equivalent. An interesting feature of this model is that it can be viewed as a difference equation expressing the expected price as a function of its own lagged value and the forcing variable p_{t-1}.

A. Find the homogeneous solution for p_t^*

Answer: Form the homogeneous equation $p_t^* - (1 - \alpha)\, p_{t-1}^* = 0$.
The homogeneous solution is:

$$\boxed{p_t^* = A(1\text{-}\alpha)^t}$$

where A is an arbitrary constant and (1-α) is the characteristic root.

B. Use lag operators to find the particular solution. Check your answer by substituting your answer into the original difference equation.

Answer: The particular solution can be written as

$$[\,1 - (1\text{-}\alpha)L\,]\ p_t^* = \alpha p_{t\text{-}1}$$

or $\quad p_t^* = \alpha p_{t\text{-}1}/[\,1 - (1\text{-}\alpha)L\,]$ so that

$$\boxed{p_t^* = \alpha[p_{t\text{-}1} + (1\text{-}\alpha)p_{t\text{-}2} + (1\text{-}\alpha)^2 p_{t\text{-}3} + \ldots\,]}$$

To check the answer, substitute the particular solution into the original difference equation

$$\alpha[p_{t\text{-}1} + (1\text{-}\alpha)p_{t\text{-}2} + (1\text{-}\alpha)^2 p_{t\text{-}3} + \ldots\,] = \alpha p_{t\text{-}1} + (1\text{-}\alpha)\alpha[p_{t\text{-}2} + (1\text{-}\alpha)p_{t\text{-}3} + (1\text{-}\alpha)^2 p_{t\text{-}4} + \ldots\,]$$

It should be clear that the equation holds as an identity.

3. Suppose that the money supply process has the form $m_t = m + \rho m_{t\text{-}1} + \varepsilon_t$ where m is a constant and $0 < \rho < 1$.

A. Show that it is possible to express m_{t+n} in terms of the known value m_t and the sequence $\{\varepsilon_{t+1}, \varepsilon_{t+2}, \ldots, \varepsilon_{t+n})$.

Answer: One method is to use forward iteration. Updating the money supply process one period yields $m_{t+1} = m + \rho m_t + \varepsilon_{t+1}$. Update again to obtain

$$\begin{aligned} m_{t+2} &= m + \rho m_{t+1} + \varepsilon_{t+2} \\ &= m + \rho[m + \rho m_t + \varepsilon_{t+1}] + \varepsilon_{t+2} = m + \rho m + \varepsilon_{t+2} + \rho\varepsilon_{t+1} + \rho^2 m_t \end{aligned}$$

Repeating the process for m_{t+3}

$$m_{t+3} = m + \rho m_{t+2} + \varepsilon_{t+3}$$

$= m + \varepsilon_{t+3} + \rho[m + \rho m + \varepsilon_{t+2} + \rho\varepsilon_{t+1} + \rho^2 m_t]$

For any period $t+n$, the solution is

$$\boxed{m_{t+n} = m(1 + \rho + \rho^2 + \rho^3 + ... + \rho^{n-1}) + \varepsilon_{t+n} + \rho\varepsilon_{t+n-1} + ... + \rho^{n-1}\varepsilon_{t+1} + \rho^n m_t}$$

B. Suppose that all values of ε_{t+i} for $i > 0$ have a mean value of zero. Explain how you could use your result in part A to forecast the money supply n-periods into the future.

Answer: The expectation of ε_{t+1} through ε_{t+n} is equal to zero. Hence, the expectation of the money supply n periods into the future is

$$\boxed{m(1 + \rho + \rho^2 + \rho^3 + ... + \rho^{n-1}) + \rho^n m_t}$$

As $n \to \infty$, the forecast approaches $m/(1-\rho)$.

4. Find the particular solutions for each of the following:

i. $y_t = a_1 y_{t-1} + \varepsilon_t + \beta_1\varepsilon_{t-1}$

Answer: Assuming $|a_1| < 1$, you can use lag operators to write the equation as $(1 - a_1L)y_t = \varepsilon_t + \beta_1\varepsilon_{t-1}$. Hence, $y_t = (\varepsilon_t + \beta_1\varepsilon_{t-1})/(1 - a_1L)$.

Now apply the expression $(1 - a_1L)^{-1}$ to each term in the numerator so that

$$y_t = \varepsilon_t + a_1\varepsilon_{t-1} + (a_1)^2\varepsilon_{t-2} + (a_1)^3\varepsilon_{t-3} + ... + \beta_1[\varepsilon_{t-1} + a_1\varepsilon_{t-2} + (a_1)^2\varepsilon_{t-3} + ...]$$

$$\boxed{y_t = \varepsilon_t + (a_1+\beta_1)\varepsilon_{t-1} + a_1(a_1+\beta_1)\varepsilon_{t-2} + (a_1)^2(a_1+\beta_1)\varepsilon_{t-3} + (a_1)^3(a_1+\beta_1)\varepsilon_{t-4} + ...}$$

If $a_1 = 1$, the improper form of the particular solution is:

$$y_t = b_0 + \varepsilon_t + (1+\beta_1)\sum_{i=1}^{\infty}\varepsilon_{t-i}$$

where: an initial condition is needed to eliminate the constant b_0 and the non-convergent sequence.

ii. $y_t = a_1 y_{t-1} + \varepsilon_{1t} + \beta\varepsilon_{2t}$

Answer: Write the equation as $y_t = \varepsilon_{1t}/(1-a_1L) + \beta\varepsilon_{2t}/(1-a_1L)$. Now, apply $(1 - a_1L)^{-1}$ to each term in the numerator so that

$y_t = \varepsilon_{1t} + a_1\varepsilon_{1t-1} + (a_1)^2\varepsilon_{1t-2} + (a_1)^3\varepsilon_{1t-3} + ... + \beta[\varepsilon_{2t} + a_1\varepsilon_{2t-1} + (a_1)^2\varepsilon_{2t-2} + (a_1)^3\varepsilon_{2t-3} + ...]$

Alternatively, you can use the Method of Undetermined Coefficients and write the challenge solution in the form

$$\boxed{\begin{array}{l} y_t = \Sigma c_i\varepsilon_{1t-i} + \Sigma d_i\varepsilon_{2t-i} \\ \textit{where} \text{ the coefficients satisfy: } c_i = (a_1)^i \text{ and } d_i = \beta(a_1)^i. \end{array}}$$

5. The *Unit Root Problem* in time-series econometrics is concerned with characteristic roots that are equal to unity. In order to preview the issue:

A. Find the homogeneous solution to each of the following.

i) $y_t = a_0 + 1.5y_{t-1} - 0.5y_{t-2} + \varepsilon_t$
Answer: The homogeneous equation is $y_t - 1.5y_{t-1} + .5y_{t-2} = 0$. The homogeneous solution will take the form $y_t = A\alpha^t$. To form the characteristic equation, first substitute this *challenge solution* into the homogeneous equation to obtain

$A\alpha^t - 1.5A\alpha^{t-1} + 0.5A\alpha^{t-2} = 0$

Next, divide by $A\alpha^{t-2}$ to obtain the characteristic equation

$\alpha^2 - 1.5\alpha + 0.5 = 0$

The two characteristic roots are $\alpha_1 = 1$, $\alpha_2 = 0.5$. The linear combination of the two homogeneous solutions is also a solution. Hence, letting A_1 and A_2 be two arbitrary constants, the complete homogeneous solution is

$$\boxed{A_1 + A_2(0.5)^t}$$

ii) $y_t = a_0 + y_{t-2} + \varepsilon_t$
Answer: The homogeneous equation is $y_t - y_{t-2} = 0$. The homogeneous solution will take the form $y_t = A\alpha^t$. To form the characteristic equation, first substitute this *challenge solution* into the homogeneous equation to obtain

$A\alpha^t - A\alpha^{t-2} = 0$

Next, divide by $A\alpha^{t-2}$ to obtain the characteristic equation $\alpha^2 - 1 = 0$. The two characteristic roots are $\alpha_1 = 1$, $\alpha_2 = -1$. The linear combination of the two homogeneous solutions is also a solution. Hence, letting A_1 and A_2 be two arbitrary constants, the complete homogeneous solution is

$$\boxed{A_1 + A_2(-1)^t}$$

iii) $y_t = a_0 + 2y_{t-1} - y_{t-2} + \varepsilon_t$

Answer: The homogeneous equation is $y_t - 2y_{t-1} + y_{t-2} = 0$. The homogeneous solution always takes the form $y_t = A\alpha^t$. To form the characteristic equation, first substitute this *challenge solution* into the homogeneous equation to obtain

$$A\alpha^t - 2A\alpha^{t-1} + A\alpha^{t-2} = 0$$

Next, divide by $A\alpha^{t-2}$ to obtain the characteristic equation

$$\alpha^2 - 2\alpha + 1 = 0$$

The two characteristic roots are $\alpha_1 = 1$, and $\alpha_2 = 1$; hence there is a repeated root. The linear combination of the two homogeneous solutions is also a solution. Letting A_1 and A_2 be two arbitrary constants, the complete homogeneous solution is

$$\boxed{A_1 + A_2 t}$$

iv) $y_t = a_0 + y_{t-1} + 0.25y_{t-2} - 0.25y_{t-3} + \varepsilon_t$

Answer: The homogeneous equation is $y_t - y_{t-1} - 0.25y_{t-2} + 0.25y_{t-3} = 0$. The homogeneous solution always takes the form $y_t = A\alpha^t$. To form the characteristic equation, first substitute this *challenge solution* into the homogeneous equation to obtain

$$A\alpha^t - A\alpha^{t-1} - 0.25A\alpha^{t-2} + 0.25A\alpha^{t-3} = 0$$

Next, divide by $A\alpha^{t-3}$ to obtain the characteristic equation

$$\alpha^3 - \alpha^2 - 0.25\alpha + 0.25 = 0$$

The three characteristic roots are $\alpha_1 = 1$, $\alpha_2 = 0.5$, and $\alpha_3 = -0.5$. The linear combination of the three homogeneous solutions is also a solution. Hence, letting A_1, A_2 and A_3 be three arbitrary constants, the complete homogeneous solution is

$$\boxed{A_1 + A_2(0.5)^t + A_3(-0.5)^t}$$

B. Show that each of the backward-looking particular solutions is not convergent.

i) $y_t = a_0 + 1.5y_{t-1} - .5y_{t-2} + \varepsilon_t$

Answer: Using lag operators, write the equation as $(1 - 1.5L + 0.5L^2)y_t = a_0 + \varepsilon_t$. Factoring the polynomial yields $(1 - L)(1 - 0.5L)y_t = a_0 + \varepsilon_t$. Although the expression $(a_0 + \varepsilon_t)/(1 - 0.5L)$ is convergent, $(a_0 + \varepsilon_t)/(1 - L)$ does not converge.

ii) $y_t = a_0 + y_{t-2} + \varepsilon_t$

Answer: Using lag operators, write the equation as $(1 - L)(1 + L)y_t = a_0 + \varepsilon_t$. It is clear that neither $(a_0 + \varepsilon_t)/(1 - L)$ nor $(a_0 + \varepsilon_t)/(1 + L)$ converges.

iii) $y_t = a_0 + 2y_{t-1} - y_{t-2} + \varepsilon_t$

Answer: Using lag operators, write the equation as $(1 - L)(1 - L)y_t = a_0 + \varepsilon_t$. Here there are two characteristic roots that equal unity. Dividing $(a_0 + \varepsilon_t)$ by either of the $(1 - L)$ expressions does not lead to a convergent result.

iv) $y_t = a_0 + y_{t-1} + 0.25y_{t-2} - 0.25y_{t-3} + \varepsilon_t$

Answer: Using lag operators, write the equation as $(1 - L)(1 - 0.5L)(1 + 0.5L)y_t = a_0 + \varepsilon_t$. The expressions $(a_0 + \varepsilon_t)/(1 + 0.5L)$ and $(a_0 + \varepsilon_t)/(1 - 0.5L)$ are convergent, but the expression $(a_0 + \varepsilon_t)/(1 - L)$ is not convergent.

C. Show that equation (i) can be written entirely in first-differences; i.e., $\Delta y_t = a_0 + .5\Delta y_{t-1} + \varepsilon_t$. Find the particular solution for Δy_t. [HINT: Define $y_t^* = \Delta y_t$ so that $y_t^* = a_0 - 0.5\, y_{t-1}^* + \varepsilon_t$. Find the particular solution for y_t^* in terms of the $\{\varepsilon_t\}$ sequence.]

Answer: Subtract y_{t-1} from each side of $y_t = a_0 + 1.5y_{t-1} - .5y_{t-2} + \varepsilon_t$ to obtain

$$
\begin{aligned}
y_t - y_{t-1} &= a_0 + 0.5y_{t-1} - .5y_{t-2} + \varepsilon_t \quad \text{so that} \\
\Delta y_t &= a_0 + 0.5y_{t-1} - 0.5y_{t-2} + \varepsilon_t \\
&= a_0 + 0.5\Delta y_{t-1} + \varepsilon_t
\end{aligned}
$$

The particular solution for $y_t^* = a_0 + 0.5\; y_{t-1}^* + \varepsilon_t$ is given by

$y_t^* = (a_0 + \varepsilon_t)/(1 - 0.5L)$ so that

$$\boxed{y_t^* = 2a_0 + \varepsilon_t + 0.5\varepsilon_{t-1} + 0.25\varepsilon_{t-2} + 0.125\varepsilon_{t-3} + \ldots.}$$

D. Similarly transform the other equations into their first-difference form. Find the backward-looking particular solution, if it exists, for the transformed equations.

ii) $y_t = a_0 + y_{t-2} + \varepsilon_t$,

Answer: Subtract y_{t-1} from each side to form $y_t - y_{t-1} = a_0 - y_{t-1} + y_{t-2} + \varepsilon_t$ or $\Delta y_t = a_0 - \Delta y_{t-1} + \varepsilon_t$ so that

$$y_t^* = a_0 - y_{t-1}^* + \varepsilon_t$$

Note that the first difference Δy_t has characteristic root that is equal to -1. The proper form of the backward-looking solution does not exist for this equation. If you attempt the challenge solution $y_t^* = b_0 + \Sigma\alpha_i\varepsilon_{t-i}$, you find

$$b_0 + \alpha_0\varepsilon_t + \alpha_1\varepsilon_{t-1} + \alpha_2\varepsilon_{t-2} + \alpha_3\varepsilon_{t-3} + ... = a_0 - b_0 - \alpha_0\varepsilon_{t-1} - \alpha_1\varepsilon_{t-2} - \alpha_2\varepsilon_{t-3} - ... + \varepsilon_t$$

Matching coefficients on like terms yields

$b_0 = a_0 - b_0$ $\qquad\Rightarrow b_0 = a_0/2$
$\alpha_0 = 1$
$\alpha_1 = -\alpha_0$ $\qquad\Rightarrow \alpha_1 = -1$
and
$\alpha_i = (-1)^i$

In Part E, students are asked to solve an equation of this form with a given initial condition.

iii) $y_t = a_0 + 2y_{t-1} - y_{t-2} + \varepsilon_t$
Answer: Subtract y_{t-1} from each side to obtain $y_t - y_{t-1} = a_0 + y_{t-1} - y_{t-2} + \varepsilon_t$ so that

$$\Delta y_t = a_0 + \Delta y_{t-1} + \varepsilon_t$$

Using the definition of y_t^* it follows that $y_t^* = a_0 + y_{t-1}^* + \varepsilon_t$. Again, a proper form for the particular solution does not exist. The improper form is

$$y_t^* = a_0 t + \varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + ...$$

Notice that the second difference $\Delta^2 y_t$ does have a convergent solution since

$$\Delta y_t^* = a_0 + \varepsilon_t$$

iv) $y_t = a_0 + y_{t-1} + 0.25y_{t-2} - 0.25y_{t-3} + \varepsilon_t$
Answer: Subtract y_{t-1} from each side and note that $0.25y_{t-2} - 0.25y_{t-3} = 0.25\Delta y_{t-2}$ so that

$\Delta y_t = a_0 + 0.25\Delta y_{t-2} + \varepsilon_t$ or
$y_t^* = a_0 + 0.25\ y_{t-2}^* + \varepsilon_t$

Write the equation as $(1 - 0.25L^2)\ y_t^* = a_0 + \varepsilon_t$. Since $(1 - 0.25L^2) = (1 - 0.5L)(1 + 0.5L)$, it follows that

$$\boxed{y_t^* = (a_0 + \varepsilon_t)/[(1 - 0.5L)(1 + 0.5L)]}$$

E. Given the initial condition y_0, find the solution for: $y_t = a_0 - y_{t-1} + \varepsilon_t$.

Answer: You can use iteration or the Method of Undetermined Coefficients to verify that the solution is

$$\boxed{y_t = \sum_{i=1}^{t} (-1)^{i+t}\, \varepsilon_i + (-1)^t\, y_0 + \frac{a_0}{2}[1 - (-1)^t\,]}$$

Using the iterative method, $y_1 = a_0 + \varepsilon_1 - y_0$ and $y_2 = a_0 + \varepsilon_2 - y_1$ so that

$y_2 = a_0 + \varepsilon_2 - a_0 - \varepsilon_1 + y_0 = \varepsilon_2 - \varepsilon_1 + y_0$

Since $y_3 = a_0 + \varepsilon_3 - y_2$, it follows that $y_3 = a_0 + \varepsilon_3 - \varepsilon_2 + \varepsilon_1 - y_0$. Continuing in this fashion yields

$y_4 = a_0 + \varepsilon_4 - y_3 = a_0 + \varepsilon_4 - a_0 - \varepsilon_3 + \varepsilon_2 - \varepsilon_1 + y_0 = \varepsilon_4 - \varepsilon_3 + \varepsilon_2 - \varepsilon_1 + y_0$

To confirm the solution for y_t note that $(-1)^{i+t}$ is positive for even values of $(i+t)$ and negative for odd values of $(i+t)$, $(-1)^t$ is positive for even values of t, and $(a_0/2)[1 - (-1)^t]$ equals zero when t is even and a_0 when t is odd.

6. A researcher estimated the following relationship for the inflation rate (π_t):

$\pi_t = -.05 + 0.7\pi_{t-1} + 0.6\pi_{t-2} + \varepsilon_t$

A. Suppose that in periods 0 and 1, the inflation rate was 10% and 11%, respectively. Find the homogeneous, particular, and general solutions for the inflation rate.

Answer: The homogeneous equation is $\pi_t - 0.7\pi_{t-1} - 0.6\pi_{t-2} = 0$. Try the challenge solution $\pi_t = A\alpha^t$, so that the characteristic equation is

$A\alpha^t - 0.7A\alpha^{t-1} - 0.6A\alpha^{t-2} = 0$ or
$\alpha^2 - 0.7\alpha - 0.6 = 0$

The characteristic roots are: $\alpha_1 = 1.2$, $\alpha_2 = -0.5$. Thus, the homogeneous solution is

$$\boxed{\pi_t = A_1(1.2)^t + A_2(-0.5)^t}$$

The backward-looking particular solution is explosive. Try the challenge solution: $\pi_t = b + \Sigma b_i\varepsilon_{t-i}$. For this to be a solution, it must satisfy

$$b + b_0\varepsilon_t + b_1\varepsilon_{t-1} + b_2\varepsilon_{t-2} + b_3\varepsilon_{t-3} + \ldots = -0.05 + 0.7(b + b_0\varepsilon_{t-1} + b_1\varepsilon_{t-2} + b_2\varepsilon_{t-3} + b_3\varepsilon_{t-4} + \ldots) + 0.6(b + b_0\varepsilon_{t-2} + b_1\varepsilon_{t-3} + b_2\varepsilon_{t-4} + b_3\varepsilon_{t-5} + \ldots) + \varepsilon_t$$

Matching coefficients on like terms yields:

$b = -0.05 + 0.7b + 0.6b \qquad \Rightarrow b = 1/6$

$b_0 = 1$

$b_1 = 0.7b_0 \qquad \Rightarrow b_1 = 0.7$

$b_2 = 0.7b_1 + 0.6b_0 \qquad \Rightarrow b_2 = 0.49 + 0.6 = 1.09$

All successive values for b_i satisfy the explosive difference equation

$$\boxed{b_i = 0.7b_{i-1} + 0.6b_{i-2}}$$

If you continue in this fashion, the successive values of the b_i are:
b_3 =1.183; b_4 = 1.4821; b_5 = 1.74727; b_6 = 2.11235; b_7 = 2.527007...

Note that the forward-looking solution is not satisfactory here unless you are willing to assume perfect foresight. However, this is inconsistent with the presence of the error term. (After all, the regression would not have to be estimated if everyone had perfect foresight.) The point is that the forward-looking solution expresses the current inflation rate in terms of the future values of the $\{\varepsilon_t\}$ sequence. If $\{\varepsilon_t\}$ is assumed to be a white-noise process, it does not make economic sense to posit that the current inflation rate is determined by the future realizations of ε_{t+i}.

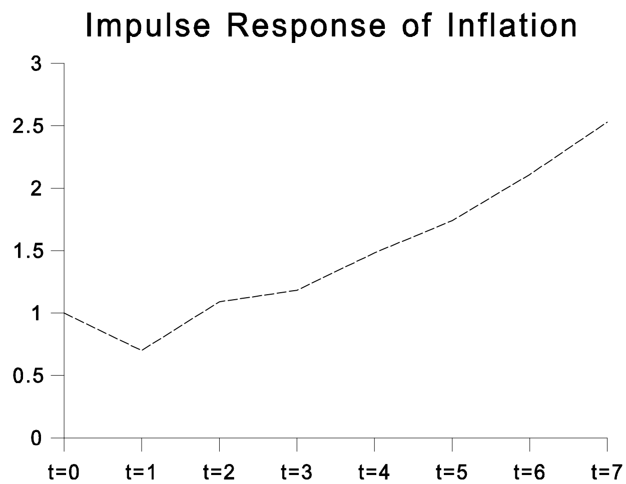
Although the backward-looking particular solution is not convergent, imposing the initial conditions on the particular solution yields finite values for all π_t (as long as t is finite). Consider the general solution

$$\pi_t = 1/6 + \varepsilon_t + 0.7\varepsilon_{t-1} + b_2\varepsilon_{t-2} + ... + b_{t-2}\varepsilon_2 + b_{t-1}\varepsilon_1 + b_t\varepsilon_0 + b_{t+1}\varepsilon_{-1} + ... + A_1(1.2)^t + A_2(-0.5)^t$$

For $t = 0$ and $t = 1$:

$$0.10 = 1/6 + \varepsilon_0 + 0.7\varepsilon_{-1} + b_2\varepsilon_{-2} + ... + A_1 + A_2$$
$$0.11 = 1/6 + \varepsilon_1 + 0.7\varepsilon_0 + b_2\varepsilon_{-1} + ... + A_1(1.2) + A_2(-0.5)$$

These last two equations define A_1 and A_2. Inserting the solutions for A_1 and A_2 into the general solution for π_t eliminates the arbitrary constants.



B. Discuss the shape of the impulse response function. Given that the U.S. is not headed for runaway inflation, why do you believe that the researcher's equation is poorly estimated?

Answer: The impulse response function is given by the $\{b_i\}$ sequence. The impact of an ε_t shock on the rate of inflation is 1. Only 70% of this initial effect remains for one period. After this one-time decay, the effect of the ε_t shock on π_{t+2}, π_{t+3}, ... explodes. You can see the impulse response function in the accompanying chart. The impulse responses imply that the inflation rate will grow exponentially. Given that there will not be runaway inflation, we would want to disregard the estimated model.

7. Consider the stochastic process: $y_t = a_0 + a_2y_{t-2} + \varepsilon_t$.

A. Find the homogeneous solution and determine the stability condition.

Answer: The homogeneous solution has the form $y_t = A\alpha^t$. Form the characteristic equation by substitution of the challenge solution into the original equation, so that

$A\alpha^t - a_2A\alpha^{t-2} = 0$ so that $\alpha^2 = a_2$.

The two characteristic roots are $\alpha_1 = \sqrt{a_2}$ and $\alpha_2 = -\sqrt{a_2}$. The stability condition is for a_2 to be less than unity in absolute value.

B. Find the particular solution using the Method of Undetermined Coefficients.

Answer: Try the challenge solution $y_t = b + \Sigma b_i\varepsilon_{t-i}$. For this to be a solution, it must satisfy

$$b + b_0\varepsilon_t + b_1\varepsilon_{t-1} + b_2\varepsilon_{t-2} + b_3\varepsilon_{t-3} + ... = a_0 + a_2(b + b_0\varepsilon_{t-2} + b_1\varepsilon_{t-3} + b_2\varepsilon_{t-4} + b_3\varepsilon_{t-5} + ...) + \varepsilon_t$$

Matching coefficients on like terms

$b = a_0 + a_2b$ $\qquad\qquad \Rightarrow b = a_0/(1-a_2)$

$b_0 = 1$
$b_1 = 0$
$b_2 = a_2 b_0 \quad \Rightarrow b_2 = a_2$
$b_3 = a_2 b_1 \quad \Rightarrow b_3 = 0$ (since $b_1 = 0$)

Continuing in this fashion, it follows that

$$\boxed{b_i = (a_2)^{i/2} \text{ if } i \text{ is even and 0 if } i \text{ is odd.}}$$

8. Consider the Cagan demand for money function in which: $m_t - p_t = \alpha - \beta[p_{t+1} - p_t]$
A. Show that the backward-looking particular solution for p_t is divergent.

Answer: Using lag operators, rewrite the equation as $\beta p_{t+1} - (1 + \beta)p_t = \alpha - m_t$. Combining terms yields $[1 - (1 + 1/\beta)L]p_{t+1} = (\alpha - m_t)/\beta$ so that lagging by one period results in

$$[1 - (1 + 1/\beta)L]p_t = (\alpha - m_{t-1})/\beta$$

Since β is assumed to be positive, the expression $(1 + 1/\beta)$ is greater than unity. Hence, the backward-looking solution for p_t is divergent.

B. Obtain the forward-looking particular solution for p_t in terms of the $\{m_t\}$ sequence. In forming the general solution, why is it necessary to assume that the money market is in long-run equilibrium?

Answer: Solving for βp_t yields

$$\beta p_t = [\alpha - m_{t-1}]/[1 - (1 + 1/\beta)L]$$

The expression $(-m_{t-1})/[1 - (1 + 1/\beta)L]$ can be obtained using Property 6 of lag operators. Let $(1 + 1/\beta)$ correspond to the term a and apply Property 6 so that

$$\begin{aligned}(-m_{t-1})/[1 - aL] &= (aL)^{-1}[(aL)^0 + (aL)^{-1} + (aL)^{-2} + (aL)^{-3} + \dots]m_{t-1} \\ &= a^{-1}[(aL)^0 + (aL)^{-1} + (aL)^{-2} + (aL)^{-3} + \dots]m_t \\ &= a^{-1}[m_t + a^{-1}m_{t+1} + a^{-2}m_{t+2} + a^{-3}m_{t+3} + \dots]\end{aligned}$$

Since $(1 + 1/\beta) = a$, it follows that $a^{-1} = \beta/(1+\beta)$. Also note that $\alpha/[1 - (1 + 1/\beta)L] = -\alpha\beta$. Thus, the solution for p_t is

$$\boxed{p_t = -\alpha + \frac{1}{1+\beta}\sum_{i=0}^{\infty}\left(\frac{\beta}{1+\beta}\right)^i m_{t+i}}$$

C. Find the impact multiplier. How does an increase in m_{t+2} affect p_t? Provide an intuitive explanation of the shape of the entire impulse response function.

Answer: The impact multiplier is the effect of m_t on p_t. As such $\partial p_t/\partial m_t = 1/(1+\beta)$. For any i, the effect of m_{t+i} on p_t is $[1/(1+\beta)][\beta/(1+\beta)]^i$. Since $[\beta/(1+\beta)]^i$ decreases as i increases, future values of the money supply have a smaller effect on the current price level than the current value. Notice that a *permanent* increase in the money supply, such that $\Delta m_t = \Delta m_{t+1} =$ ... has a proportional effect on p_t since $[1/(1+\beta)]\Sigma[(\beta/(1+\beta)]^i = 1$.

9. For each of the following, verify that the posited solution satisfies the difference equation. The symbols c, c_0, and a_0 denote constants.

Equation	Solution
A. $y_t - y_{t-1} = 0$	$y_t = c$
B. $y_t - y_{t-1} = a_0$	$y_t = c + a_0 t$
C. $y_t - y_{t-2} = 0$	$y_t = c + c_0(-1)^t$
D. $y_t - y_{t-2} = \varepsilon_t$	$y_t = c + c_0(-1)^t + \varepsilon_t + \varepsilon_{t-2} + \varepsilon_{t-4} + ...$

Answer: Substitute each posited solution into the original difference.

A. Since $y_t = c$ and $y_{t-1} = c$, it immediately follows that $c - c = 0$.

B. Since $y_{t-1} = c + a_0(t-1)$, it follows that $c + a_0 t - c - a_0(t-1) = a_0$.

C. The issue is whether $c + c_0(-1)^t - c - c_0(-1)^{t-2} = 0$? Since $(-1)^t = (-1)^{t-2}$, the posited solution is correct.

D. Does $c + c_0(-1)^t + \varepsilon_t + \varepsilon_{t-2} + \varepsilon_{t-4} + ... - c - c_0(-1)^{t-2} - \varepsilon_{t-2} - \varepsilon_{t-4} - \varepsilon_{t-6} - ... = \varepsilon_t$? Since $c_0(-1)^t = c_0(-1)^{t-2}$, the posited solution is correct.

10. Part 1: For each of the following, determine whether $\{y_t\}$ represents a stable process. Determine whether the characteristic roots are real or imaginary and whether the real parts are positive or negative.

A. $y_t - 1.2y_{t-1} + .2y_{t-2}$

B. $y_t - 1.2y_{t-1} + .4y_{t-2}$

C. $y_t - 1.2y_{t-1} - 1.2y_{t-2}$

D. $y_t + 1.2y_{t-1}$

E. $y_t - 0.7y_{t-1} - 0.25y_{t-2} + 0.175y_{t-3} = 0$ [Hint: $(x - 0.5)(x + 0.5)(x - 0.7) = x^3 - 0.7x^2 - 0.25x + 0.175$]

Answers:

A. The characteristic equation $\alpha^2 - 1.2\alpha + 0.2 = 0$ has roots $\alpha_1 = 1$ and $\alpha_2 = 0.2$. The unit root means that the $\{y_t\}$ sequence is not convergent.

B. The characteristic equation $\alpha^2 - 1.2\alpha + 0.4 = 0$ has roots $\alpha_1, \alpha_2 = 0.6 \pm 0.2i$. The roots are imaginary. The $\{y_t\}$ sequence exhibits damped wave-like oscillations.

C. The characteristic equation $\alpha^2 - 1.2\alpha - 1.2 = 0$ has roots $\alpha_1 = 1.85$ and $\alpha_2 = -0.65$. One of the roots is outside the unit circle so that the $\{y_t\}$ sequence is explosive.

D. The characteristic equation $\alpha + 1.2 = 0$ has the root $\alpha = -1.2$. The $\{y_t\}$ sequence has explosive oscillations.

E. The characteristic equation $\alpha^3 - 0.7\alpha^2 - 0.25\alpha + 0.175 = 0$ has roots $\alpha_1 = 0.7$, $\alpha_2 = 0.5$ and $\alpha_3 = -0.5$. Although all roots are real, there are damped oscillations due to the presence of the term $(-0.5)^t$.

Part 2: Write each of the above equations using lag operators. Determine the characteristic roots of the inverse characteristic equation.

> **Answers**: Rewrite each using lag operators in order to obtain the inverse characteristic equation.
>
> A. $(1 - 1.2L + 0.2L^2)y_t$ has the inverse characteristic equation $1 - 1.2L + 0.2L^2 = 0$. Solving this quadratic equation for the two values of L (called L_1 and L_2) yield the characteristic roots of the *inverse characteristic equation*. Here, $L_1 = 1.0$ and $L_2 = 5.0$. Since one root lies on the unit circle, the $\{y_t\}$ sequence is not convergent. Note that these roots are the reciprocals of the roots found in Part 1.
>
> B. $(1 - 1.2L + 0.4L^2)y_t$ has the inverse characteristic equation $1 - 1.2L + 0.4L^2 = 0$. The roots are $L_1, L_2 = 1.5 \pm 0.5i$. The roots of the inverse characteristic equation are *outside* the unit circle so that the $\{y_t\}$ sequence exhibits convergent wave-like oscillations.
>
> C. $(1 - 1.2L - 1.2L^2)y_t$ has the inverse characteristic equation $1 - 1.2L - 1.2L^2 = 0$. The roots are -1.54 and 0.54. One of the *inverse* characteristic roots is *inside* the unit circle so that the $\{y_t\}$ sequence is explosive.
>
> D. The inverse characteristic equation $(1 + 1.2L)y_t$ has the inverse characteristic root: $L = -1/1.2 = -0.8333$. Since this *inverse* characteristic root is negative and lies *inside* the unit circle, the $\{y_t\}$ sequence has explosive oscillations.
>
> E. $(1 - 0.7L - 0.25L^2 + 0.175L^3)y_t$ has the inverse characteristic equation $1 - 0.7L - 0.25L^2 + 0.175L^3 = 0$. Factoring yields the equivalent representation $(1 - 0.5L)(1 + 0.5L)(1 - 0.7L) = 0$. The inverse characteristic roots are 2.0, -2.0, and 1.0/0.7 = 1.429. All the inverse characteristic roots lie outside of the unit circle.

11. Consider the stochastic difference equation: $y_t = 0.8y_{t-1} + \varepsilon_t - 0.5\varepsilon_{t-1}$.

A. Suppose that the initial conditions are such that: $y_0 = 0$ and $\varepsilon_0 = \varepsilon_{-1} = 0$. Now suppose that $\varepsilon_1 = 1$. Determine the values y_1 through y_5 by forward iteration.

> **Answer:** If we assume that all future values of $\{\varepsilon_t\} = 0$ we can find the solution. In essence, this is the method used to obtain the impulse response function.

$$\boxed{y_1 = 1, y_2 = 0.3, y_3 = 0.24, y_4 = 0.192, y_5 = 0.1536}$$

B. Find the homogeneous and particular solutions.

> **Answer**: The solution to the homogeneous equation $y_t - 0.8y_{t-1} = 0$ is $y_t = A(0.8)^t$.
>
> Using lag operators, the particular solution is $y_t = (\varepsilon_t - 0.5\varepsilon_{t-1})/(1 - 0.8L)$. If we apply $1/(1-0.8L)$ to ε_t and $-0.5\varepsilon_{t-1}$, we obtain

$$
\begin{aligned}
y_t &= \varepsilon_t + 0.8\varepsilon_{t-1} + (0.8)^2\varepsilon_{t-2} + (0.8)^3\varepsilon_{t-3} + \ldots -0.5[\varepsilon_{t-1} + 0.8\varepsilon_{t-2} + (0.8)^2\varepsilon_{t-3} + \ldots] \\
&= \varepsilon_t + (0.8 - 0.5)\varepsilon_{t-1} + 0.8(0.8 - 0.5)\varepsilon_{t-2} + 0.8^2(0.8 - 0.5)\varepsilon_{t-3} + \ldots
\end{aligned}
$$

$$\boxed{y_t = \varepsilon_t + 0.3\varepsilon_{t-1} + 0.8(0.3)\varepsilon_{t-2} + 0.8^2(0.3)\varepsilon_{t-3} + \ldots}$$

C. Impose the initial conditions in order to obtain the general solution.

Answer: Combining the homogeneous and particular solutions yields the general solution

$y_t = \varepsilon_t + 0.3\varepsilon_{t-1} + 0.8(0.3)\varepsilon_{t-2} + 0.8^2(0.3)\varepsilon_{t-3} + \ldots + A(0.8)^t$.

Now impose the initial condition $y_0 = 0$ and $\varepsilon_0 = \varepsilon_{-1} = 0$ to obtain

$0 = \varepsilon_0 + 0.3\varepsilon_{-1} + 0.8(0.3)\varepsilon_{-2} + 0.8^2(0.3)\varepsilon_{-3} + \ldots + A$. Hence
$A = -\varepsilon_0 - 0.3\varepsilon_{-1} - 0.8(0.3)\varepsilon_{-2} - 0.8^2(0.3)\varepsilon_{-3} + \ldots$

Hence, $A = 0$ if the system began in initial equilibrium. Now substitute for A to obtain

$$\boxed{y_t = \varepsilon_t + 0.3\sum_{i=0}^{t-2}(0.8)^i\,\varepsilon_{t-i-1}}$$

D. Trace out the time path of an ε_t shock on the entire time path of the $\{y_t\}$ sequence.

Answer: $\partial y_t/\partial\varepsilon_t = 1$; $\partial y_{t+1}/\partial\varepsilon_t = \partial y_t/\partial\varepsilon_{t-1} = 0.3$; $\partial y_{t+2}/\partial\varepsilon_t = \partial y_t/\partial\varepsilon_{t-2} = 0.3(0.8)$; $\partial y_{t+3}/\partial\varepsilon_t = \partial y_t/\partial\varepsilon_{t-3} = 0.3(0.8)^2$; and for $i \geq 1$:

$$\boxed{\partial y_{t+i}/\partial\varepsilon_t = \partial y_t/\partial\varepsilon_{t-i} = 0.3(0.8)^{i-1}}$$

12. Use Equation (1.5) to determine the restrictions on α and β necessary to ensure that the $\{y_t\}$ process is stable.

Answer: To determine stability, it is only necessary to examine the homogeneous portion of (1.5); i.e., $y_t - \alpha(1+\beta)y_{t-1} + \alpha\beta y_{t-2} = 0$ where $0 < \alpha < 1$ and $\beta > 0$.

In terms of the notation used in Figure 1.6, $a_1 = \alpha(1+\beta)$ and $a_2 = -\alpha\beta$. Given that α and β are positive, $a_1 > 0$ and $a_2 < 0$. Thus, the point labled α_2 could correspond to $\alpha(1+\beta)$ units along the a_1 axis and $-\alpha\beta$ units along the a_2 axis. The stability conditions for a second-order difference equation are:

$a_1 + a_2 < 1$
$a_2 < 1 + a_1$
$-a_2 < 1$ (since $a_2 < 0$).

Note that $a_1 + a_2 = \alpha(1+\beta) - \alpha\beta = \alpha$. Since $0 < \alpha < 1$, the first stability condition is always satisfied. To satisfy the second condition (i.e., $a_2 < 1 + a_1$), it is necessary to restrict the coefficients such that $-\alpha\beta < 1 + \alpha(1+\beta)$; simple manipulation yields: $0 < 1 + \alpha + 2\alpha\beta$. Since α and β are positive, the second stability condition necessarily holds. The third condition (i.e., $-a_2 < 1$) is equivalent to $\alpha\beta < 1$ or $\beta < 1/\alpha$. Hence, to ensure stability, it is necessary to restrict β to be less than $1/\alpha$.

CHAPTER 2

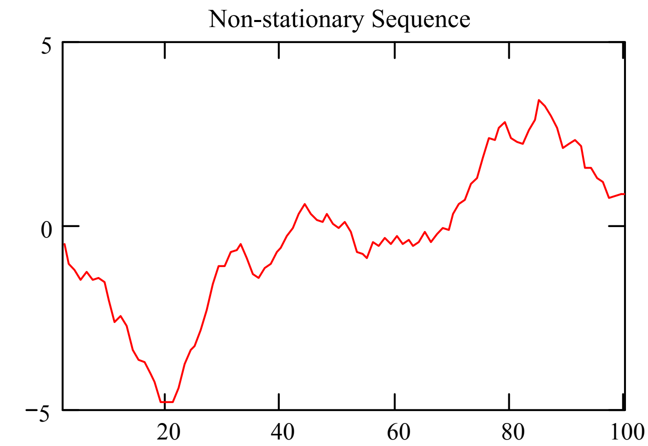
STATIONARY TIME-SERIES MODELS

Lecture Suggestions

I use Figure M2-1 to illustrate the effects of differencing and over-differencing. The first graph depicts 100 realizations of the unit root process $y_t = 1.5y_{t-1} - 0.5y_{t-2} + \varepsilon_t$. If you examine the graph, it is clear there is no tendency for mean reversion. This non-stationary series has a unit root that can be eliminated by differencing. The second graph in the figure shows the first-difference of the $\{y_t\}$ sequence: $\Delta y_t = 0.5\Delta y_{t-1} + \varepsilon_t$. The positive autocorrelation ($\rho_1 = 0.5$) is reflected in the tendency for large (small) values of Δy_t to be followed by other large (small) values. It is simple to make the point that the $\{\Delta y_t\}$ sequence can be estimated using the Box-Jenkins methodology. It is obvious to students that the ACF will reflect the positive autocorrelation. The third graph shows the second difference: $\Delta^2 y_t = -0.5\Delta^2 y_{t-1} + \varepsilon_t - \varepsilon_{t-1}$. Students are quick to understand the difficulties of estimating this over-differenced series. Due to the extreme volatility of the $\{\Delta^2 y_t\}$ series, the current value of $\Delta^2 y_t$ is not helpful in predicting $\Delta^2 y_{t+1}$.

The effects of logarithmic data transformations are often taken for granted. I use Figure M2-2 to illustrate the effects of the Box-Cox transformation. The first graph shows 100 realizations of the simulated AR(1) process: $y_t = 5 + 0.5y_{t-1} + \varepsilon_t$. The $\{\varepsilon_t\}$ series is precisely the same as that used in constructing the graphs in Figure M2-1. In fact, the only difference between

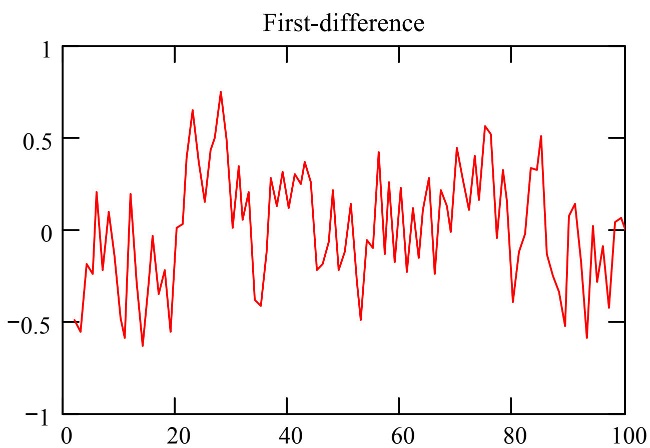
Figure M2-1: The Effects of Differencing



The $\{y_t\}$ sequence was constructed as:

$$y_t = 1.5y_{t-1} - 0.5y_{t-2} + \varepsilon_t$$

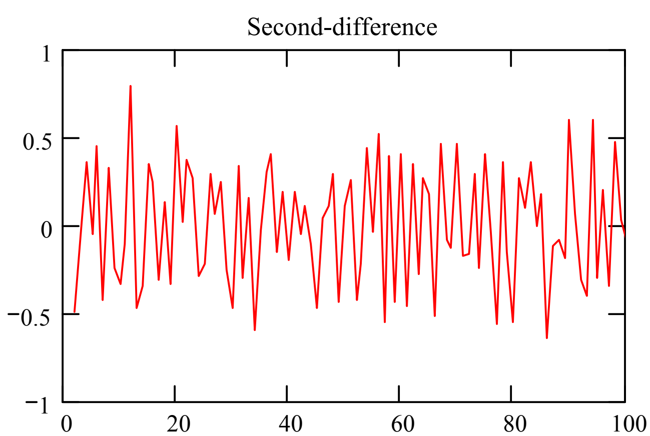
The unit root means that the sequence does not exhibit any tendency for mean reversion.



The first-difference of the $\{y_t\}$ sequence is:

$$\Delta y_t = 0.5\Delta y_{t-1} + \varepsilon_t$$

The first-difference of $\{\Delta y_t\}$ is a stationary AR(1) process such that $a_1 = 0.5$.

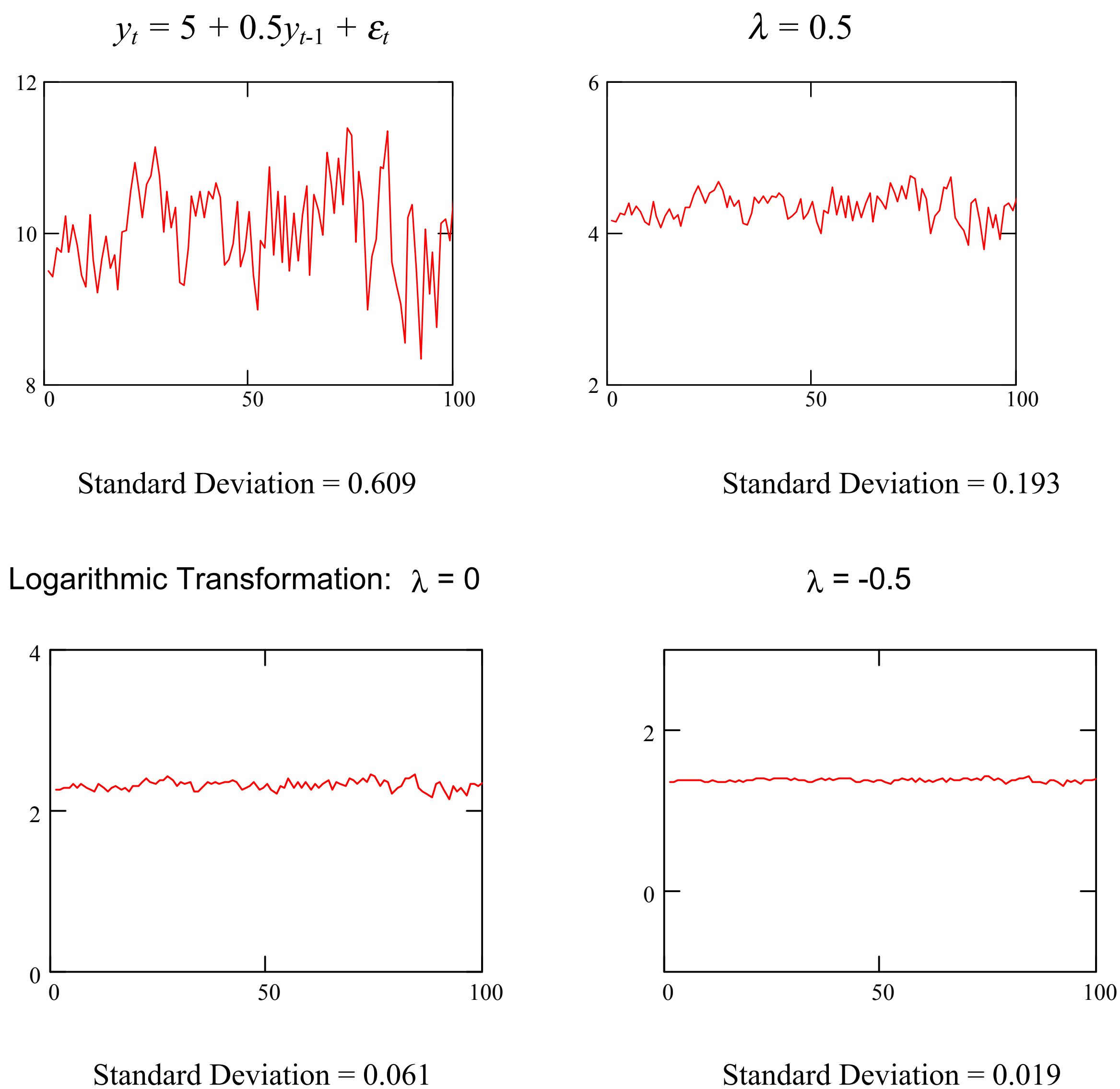


The second-difference of the sequence is:

$$\Delta^2 y_t = -0.5\Delta^2 y_{t-1} + \varepsilon_t - \varepsilon_{t-1}$$

The over-differenced $\{\Delta^2 y_t\}$ sequence has an invertible error term.

Figure M2-2: Box-Cox Transformations



The first graph shows 100 realizations of a simulated AR(1) process; by construction, the standard deviation of the $\{y_t\}$ sequence is 0.609. The next three graphs show the results of Box-Cox transformations using values of $\lambda = 0.5$, 0.0, and -0.5, respectively. You can see that decreasing λ acts to smooth the sequence.

the middle graph of Figure M2-1 and the first graph of Figure M2-2 involves the presence of the intercept term. The effects of a logarithmic transformation can be seen by comparing the two left-hand-side graphs of Figure M2-2. It should be clear that the logarithmic transformation smooths" the series. The natural tendency is for students to think smoothing is desirable. However, I point out that actual data (such as asset prices) can be quite volatile and that individuals may respond to the volatility of the data and not the logarithm of the data. Thus, there may be instances in which we do not want to reduce the variance actually present in the data. At this time, I mention that the material in Chapter 3 shows how to estimate the conditional variance of a series. Two Box-Cox transformations are shown in the right-hand-side graphs of Figure M2-2. Notice that decreasing λ reduces variability and that a small change in λ can have a pronounced effect on the variance.

Answers to Questions

1. In the coin-tossing example of Section 1, your winnings on the last four tosses (w_t) can be denoted by

$$w_t = 1/4\varepsilon_t + 1/4\varepsilon_{t-1} + 1/4\varepsilon_{t-2} + 1/4\varepsilon_{t-3}$$

A. Find the expected value of your winnings. Find the expected value given that $\varepsilon_{t-3} = \varepsilon_{t-2} = 1$.

Answers: Throughout the text, the term ε_t denotes a white-noise disturbance. The properties of the $\{\varepsilon_t\}$ sequence are such that:
i. $E\varepsilon_t = E\varepsilon_{t-1} = E\varepsilon_{t-2} = ... = 0$, ii. $E\varepsilon_t\varepsilon_{t-i} = 0$ for $i \neq 0$, and iii. $E(\varepsilon_t)^2 = E(\varepsilon_{t-i})^2 = ... = \sigma^2$.
Hence:

$$\boxed{Ew_t = E(1/4\varepsilon_t + 1/4\varepsilon_{t-1} + 1/4\varepsilon_{t-2} + 1/4\varepsilon_{t-3})}$$

Since the expectation of a sum is the sum of the expectations, it follows that

$$Ew_t = (1/4)(E\varepsilon_t + E\varepsilon_{t-1} + E\varepsilon_{t-2} + E\varepsilon_{t-3}) = 0.$$

Given the information $\varepsilon_{t-3} = \varepsilon_{t-2} = 1$, the conditional expectation of w_t is $E_{t-2}w_t = E(w_t \mid \varepsilon_{t-3} = \varepsilon_{t-2} = 1) = (1/4)(E_{t-2}\varepsilon_t + E_{t-2}\varepsilon_{t-1} + E_{t-2}\varepsilon_{t-2} + E_{t-2}\varepsilon_{t-3})$ so that

$$\boxed{E_{t-2}w_t \quad = 0.25(0 + 0 + 1 + 1) = 0.5}$$

B. Find var(w_t). Find var(w_t) conditional on $\varepsilon_{t-3} = \varepsilon_{t-2} = 1$.

Answers: $\text{var}(w_t) = E(w_t)^2 - [E(w_t)]^2$ so that

$$\begin{aligned}\text{var}(w_t) &= E(1/4\varepsilon_t + 1/4\varepsilon_{t-1} + 1/4\varepsilon_{t-2} + 1/4\varepsilon_{t-3})^2 \\ &= (1/16)E[(\varepsilon_t)^2 + 2\varepsilon_t\varepsilon_{t-1} + 2\,\varepsilon_t\varepsilon_{t-2} + 2\varepsilon_t\varepsilon_{t-3} + (\varepsilon_{t-1})^2 + 2\varepsilon_{t-1}\varepsilon_{t-2} + 2\varepsilon_{t-1}\varepsilon_{t-3} \\ &\qquad + (\varepsilon_{t-2})^2 + 2\varepsilon_{t-2}\varepsilon_{t-3} + (\varepsilon_{t-3})^2]\end{aligned}$$

Since the expected values of all cross-products are zero, and it follows that:

$$\boxed{\text{var}(w_t) = (1/16)4\sigma^2 = 0.25\sigma^2}$$

Given the information $\varepsilon_{t-3} = \varepsilon_{t-2} = 1$, the conditional variance is

$$\begin{aligned}\text{var}(w_t \mid \varepsilon_{t-3} = \varepsilon_{t-2} = 1) &= E_{t-2}(1/4\varepsilon_t + 1/4\varepsilon_{t-1} + 1/4\varepsilon_{t-2} + 1/4\varepsilon_{t-3})^2 - (E_{t-2}w_t)^2 \\ &= (1/16)E_{t-2}[\varepsilon_t^2 + 2\varepsilon_t\varepsilon_{t-1} + 2\varepsilon_t\varepsilon_{t-2} + 2\varepsilon_t\varepsilon_{t-3} + (\varepsilon_{t-1})^2 + 2\varepsilon_{t-1}\varepsilon_{t-2} + 2\varepsilon_{t-1}\varepsilon_{t-3} \\ &\quad + (\varepsilon_{t-2})^2 + 2\varepsilon_{t-2}\varepsilon_{t-3} + (\varepsilon_{t-3})^2] - (0.5)^2\end{aligned}$$

Since $E_{t-2}\varepsilon_{t-2} = E_{t-2}\varepsilon_{t-3} = 1$, it follows that

$\text{var}(w_t \mid \varepsilon_{t-3} = \varepsilon_{t-2} = 1) = (1/16)(\sigma^2 + \sigma^2 + 1 + 1 + 2) - 0.25$, so that

$$\boxed{\text{var}(w_t \mid \varepsilon_{t-3} = \varepsilon_{t-2} = 1) = (1/8)\sigma^2}$$

C. Find: i. $\text{Cov}(w_t, w_{t-1})$ ii. $\text{Cov}(w_t, w_{t-2})$ iii. $\text{Cov}(w_t, w_{t-5})$

Answers: Using the same techniques as in Part B:

$$\begin{aligned}\text{i. } \text{Cov}(w_t, w_{t-1}) &= Ew_tw_{t-1} - E(w_t)E(w_{t-1}) = (1/16)E(\varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3})(\varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3} + \varepsilon_{t-4}) \\ &= (1/16)E[(\varepsilon_{t-1})^2 + (\varepsilon_{t-2})^2 + (\varepsilon_{t-3})^2 + \text{cross-product terms}]\end{aligned}$$

Since the expected values of the cross-product terms are all zero

$$\boxed{\text{Cov}(w_t, w_{t-1}) = (1/16)3\sigma^2}$$

$$\begin{aligned}\text{ii. } \text{Cov}(w_t, w_{t-2}) &= (1/16)E(\varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3})(\varepsilon_{t-2} + \varepsilon_{t-3} + \varepsilon_{t-4} + \varepsilon_{t-5}) \\ &= (1/16)E[(\varepsilon_{t-2})^2 + (\varepsilon_{t-3})^2 + \text{cross-product terms}]\end{aligned}$$

Since the expected values of the cross-product terms are all zero

$$\boxed{\text{Cov}(w_t, w_{t-2}) = (1/16)2\sigma^2}$$

$$\begin{aligned}\text{iii. } \text{Cov}(w_t, w_{t-5}) &= (1/16)E(\varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3})(\varepsilon_{t-5} + \varepsilon_{t-6} + \varepsilon_{t-7} + \varepsilon_{t-8}) \\ &= (1/16)E[\text{cross-product terms}]. \text{ Hence:}\end{aligned}$$

$$\boxed{\text{Cov}(w_t, w_{t-5}) = 0}$$

2. Substitute (2.10) into $y_t = a_0 + a_1y_{t-1} + \varepsilon_t$. Show that the resulting equation is an identity.

Answer: For (2.10) to be a solution, it must satisfy:

$$\begin{aligned}&a_0[1 + a_1 + a_1^2 + \ldots + a_1^{t-1}] + a_1^ty_0 + \varepsilon_t + a_1\varepsilon_{t-1} + a_1^2\varepsilon_{t-2} + \ldots + a_1^{t-1}\varepsilon_1 = \\ &\quad a_0 + a_1\{a_0[1 + a_1 + a_1^2 + \ldots + a_1^{t-2}] + a_1^{t-1}y_0 + \varepsilon_{t-1} + a_1\varepsilon_{t-2} + a_1^2\varepsilon_{t-3} + \ldots + a_1^{t-2}\varepsilon_1\} + \varepsilon_t\end{aligned}$$

Notice that all terms cancel. Specifically:

$a_0[1 + a_1 + a_1^2 + \ldots + a_1^{t-1}] \equiv a_0 + a_1\{a_0[1 + a_1 + a_1^2 + \ldots + a_1^{t-2}]\}$

$a_1^ty_0 \equiv a_1a_1^{t-1}y_0$ and:

$\varepsilon_t + a_1\varepsilon_{t-1} + a_1^2\varepsilon_{t-2} + \ ... + a_1^{t-1}\varepsilon_1 = a_1\{\ \varepsilon_{t-1} + a_1\varepsilon_{t-2} + a_1^2\varepsilon_{t-3} + \ ... + a_1^{t-2}\varepsilon_1\} + \varepsilon_t$

A. Find the homogeneous solution to: $y_t = a_0 + a_1y_{t-1} + \varepsilon_t$.

Answer: Attempt a challenge solution of the form $y_t = A\alpha^t$. For this solution to solve the homogeneous equation it follows that $\alpha = a_1$ and A can be any arbitrary constant.

B. Find the particular solution given that $|\ a_1\ | < 1$.

Answer: Using lag operators, write the equation as $(1 - a_1L)y_t = a_0 + \varepsilon_t$. Since $a_0/(1-a_1L) = a_0/(1-a_1)$ and $\varepsilon_t/(1-a_1L) = \varepsilon_t + a_1\varepsilon_{t-1} + a_1^2\varepsilon_{t-2} + \ ... + a_1^{t-1}\varepsilon_1 + a_1^t\varepsilon_0 + a_1^{t+1}\varepsilon_{-1} + ...$, it follows that the particular solution is

$$\boxed{y_t = a_0 / (1 - a_1) + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i}}$$

C. Show how to obtain (2.10) by combining the homogeneous and particular solutions.

Answer: Combining the homogeneous and particular solutions yields the general solution:

$$y_t = a_0 / (1 - a_1) + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i} + A(a_1)^t$$

so that when $t = 0$

$$y_0 = a_0 / (1 - a_1) + \sum_{i=0}^{\infty} a_1^i \varepsilon_{-i} + A$$

Solve for A and substitute the answer into the general solution to obtain (2.10).

3. Consider the second-order autoregressive process $y_t = a_0 + a_2y_{t-2} + \varepsilon_t$,where $|a_2| < 1$.

A. Find: i. $E_{t-2}y_t$ ii. $E_{t-1}y_t$ iii. E_ty_{t+2}

iv. $\mathrm{Cov}(y_t, y_{t-1})$ v. $\mathrm{Cov}(y_t, y_{t-2})$ vi. the partial autocorrelations ϕ_{11} and ϕ_{22}

Answers:

$$\boxed{\text{i) } E_{t-2}y_t = E_{t-2}(a_0 + a_2y_{t-2} + \varepsilon_t) = a_0 + a_2y_{t-2}}$$

$$\boxed{\text{ii) } E_{t-1}y_t = E_{t-1}(a_0 + a_2y_{t-2} + \varepsilon_t) = a_0 + a_2y_{t-2}}$$

Note the $E_{t-1}y_t = E_{t-2}y_t$ since information obtained in period $(t$-1) does not help to predict the value of y_t.

iii) E_ty_{t+2} can be obtained directly from the answer to Part i. Simply update the time index by two periods to obtain:

$$\boxed{E_ty_{t+2} = a_0 + a_2y_t}$$

The simplest way to answer Parts iv. and v. is to obtain the particular solution for y_t. Students should be able to show:

$$y_t = a_0/(1-a_2) + \varepsilon_t + a_2\varepsilon_{t-2} + (a_2)^2\varepsilon_{t-4} + (a_2)^3\varepsilon_{t-6} + (a_2)^4\varepsilon_{t-8} + ...$$

iv) $\text{Cov}(y_t, y_{t-1}) = E[(y_t - Ey_t)(y_{t-1} - Ey_{t-1})]$

$$= E[\varepsilon_t + a_2\varepsilon_{t-2} + (a_2)^2\varepsilon_{t-4} + (a_2)^3\varepsilon_{t-6} + (a_2)^4\varepsilon_{t-8} + ...]$$
$$[\varepsilon_{t-1} + a_2\varepsilon_{t-3} + (a_2)^2\varepsilon_{t-5} + (a_2)^3\varepsilon_{t-7} + ...]$$

so that

$$\boxed{\text{Cov}(y_t, y_{t-1}) = 0}$$

v) $\text{Cov}(y_t, y_{t-2}) = E[\varepsilon_t + a_2\varepsilon_{t-2} + (a_2)^2\varepsilon_{t-4} + (a_2)^3\varepsilon_{t-6} + \ ...]$

$$[\varepsilon_{t-2} + a_2\varepsilon_{t-4} + (a_2)^2\varepsilon_{t-6} + (a_2)^3\varepsilon_{t-8} + ...]$$
$$= a_2E[(\varepsilon_{t-2})^2 + (a_2)^2(\varepsilon_{t-4})^2 + (a_2)^4(\varepsilon_{t-6})^2 + (a_2)^6(\varepsilon_{t-8})^2 + ...]$$

Given $|a_2| < 1$, the infinite summation $\Sigma(a_2)^{2i} = 1/[1 - (a_2)^2]$ so that:

$$\boxed{\text{cov}(y_t, y_{t-2}) = a_2\sigma^2/[1 - (a_2)^2]}$$

vi). As shown in Part iv, the covariance between y_t and y_{t-1} is zero. Hence, from (2.35), $\rho_1 = \phi_{11} = 0$. Given $\rho_1 = 0$, (2.36) indicates that $\phi_{22} = \rho_2$. Given the answer to v and that $\text{var}(y_t) = \text{var}\ (y_{t-i}) = ... = \sigma^2/[1 - (a_2)^2]$, it follows that

$$\boxed{\phi_{22} = \text{cov}(y_t, y_{t-2})/\text{var}(y_t) = \{a_2\sigma^2/[1 - (a_2)^2]\}/\{\sigma^2/[1 - (a_2)^2]\} = a_2}$$

B. Find the impulse response function. Given y_{t-2}, trace out the effects on an ε_t shock on the $\{y_t\}$ sequence.

Answer: One way to answer the question is to use the particular solution for y_t:

$$y_t = a_0/(1-a_2) + \varepsilon_t + a_2\varepsilon_{t-2} + (a_2)^2\varepsilon_{t-4} + (a_2)^3\varepsilon_{t-6} + (a_2)^4\varepsilon_{t-8} + ...$$

Hence: $\partial y_t/\partial\varepsilon_t = 1$. By the simple change of subscripts:

$$\boxed{\begin{array}{l}\partial y_{t+1}/\partial\varepsilon_t = \partial y_t/\partial\varepsilon_{t-1} = 0;\ \partial y_{t+2}/\partial\varepsilon_t = \partial y_t/\partial\varepsilon_{t-2} = a_2; \\ \partial y_{t+3}/\partial\varepsilon_t = \partial y_t/\partial\varepsilon_{t-3} = 0;\ \partial y_{t+4}/\partial\varepsilon_t = \partial y_t/\partial\varepsilon_{t-4} = (a_2)^2\ ...\end{array}}$$

C. Determine the forecast function: E_ty_{t+s}. The forecast error $e_t(s)$ is the difference between y_{t+s} and E_ty_{t+s}. Derive the correlogram of the $\{ e_t(s) \}$ sequence. [Hint: Find $E_te_t(s)$, $\text{Var}[e_t(s)]$, and $E_t[e_t(s)e_t(s-j)]$ for $j = 0$ to s].

Answer: To find the forecast function, first find the general solution for y_t in terms of y_0. For a given value of y_0, students should be able to show that if t is even, then:

$$y_t = a_0[1 + a_2 + (a_2)^2 + ... + (a_2)^{t/2-1}] + \varepsilon_t + a_2\varepsilon_{t-2} + (a_2)^2\varepsilon_{t-4} + (a_2)^3\varepsilon_{t-6} + ... + (a_2)^{t/2}y_0$$

Hence: $E_0y_t = a_0[1 + a_2 + (a_2)^2 + ... + (a_2)^{t/2-1}] + (a_2)^{t/2}y_0$. To find E_ty_{t+2s}, update the time subscripts such that

$$\boxed{E_ty_{t+2s} = a_0[1 + a_2 + (a_2)^2 + ... + (a_2)^{s-1}] + (a_2)^s y_t}$$

Generalizing the result from Part A above, note that for odd-period forecasts, $E_ty_{t+2s-1} = E_ty_{t+2s-2}$. To find the forecast error, subtract E_ty_{t+s} from y_{t+s}. For even values of s, we can let $s = j/2$ and form

$$e_t(s) = \varepsilon_{t+s} + a_2\varepsilon_{t+s-2} + (a_2)^2\varepsilon_{t+s-4} + ... + (a_2)^{s/2-1}\varepsilon_{t+2}$$

The forecast error has a mean of zero since:

$$E_t[\varepsilon_{t+s} + a_2\varepsilon_{t+s-2} + (a_2)^2\varepsilon_{t+s-4} + ... + (a_2)^{s/2-1}\varepsilon_{t+2}] = 0$$

Similarly, the variance is $[1 + (a_2)^2 + (a_2)^4 + ... + (a_2)^{s-2}]\sigma^2$.

The correlations between the forecast error for any period $t+s$ and the forecast error for any odd period is zero. For even periods, we can begin by forming the covariance between $e_t(2)$ and $e_t(4)$ as:

$E[(\varepsilon_{t+4} + a_2\varepsilon_{t+2})(\varepsilon_{t+2})] = a_2\sigma^2$ so that the correlation coefficient is $a_2/(1 + (a_2)^2)^{1/2}$.

Similarly, the covariance between $e_t(2)$ and $e_t(6)$ as:

$$E[(\varepsilon_{t+6} + a_2\varepsilon_{t+4} + (a_2)^2\varepsilon_{t+2})(\varepsilon_{t+2})] = (a_2)^2\sigma^2$$

so that the correlation coefficient is

$$\boxed{(a_2)^2/[1 + (a_2)^2 + (a_2)^4]^{1/2}}$$

4. Two different balls are drawn from a jar containing three balls numbered 1, 2, and 4. Let x = number on the first ball drawn and y = sum of the two balls drawn.

A. Find the joint probability distribution for x and y; that is, find prob(x = 1, y = 3), prob(x = 1, y = 5), ... , and prob(x = 4, y = 6).

Answer: Let x = number on the first ball; z = number on the second ball; and $y = x + z$. Consider the contingency table

Summations

	$x = 1$	$x = 2$	$x = 4$
$z = 1$		$y = 3$	$y = 5$
$z = 2$	$y = 3$		$y = 6$
$z = 4$	$y = 5$	$y = 6$	

Notice that the same outcome for x and z is not possible since two *different* balls are drawn from the jar. Thus, prob(x=1, y=2), prob(x=2, y=4), and prob(x=3, y=6) are all equal to zero. The remaining six outcomes are equally likely. The probabilities (x=1, y=3), (x=1, y=5), ... (x=4, y=6) all equal 1/6.

B. Find each of the following: $E(x)$, $E(y)$, $E(y \mid x = 1)$, $E(x \mid y = 5)$, $\text{Var}(x \mid y = 5)$, and $E(y^2)$

Answers: Each outcome for x has a probability of 1/3. Thus:

$$\boxed{\text{i) } E(x) = (1/3)(1 + 2 + 4) = 7/3}$$

ii) The expected value of y is the summation of each possible outcome for y multiplied by the probability of that outcome. Since each cell has a probability of 1/6, reading across the rows of the table yields:

$$\boxed{E(y) = (1/6)(3 + 5 + 3 + 6 + 5 + 6) = 14/3}$$

iii) When $x = 1$, y can take on values 3 or 5. Hence:

$$\boxed{E(y \mid x = 1) = 0.5(3) + 0.5(5) = 4.0}$$

iv) When $y = 5$, x can take on values 1 or 4. Hence:

$$\boxed{E(x \mid y = 5) = 0.5(1 + 4) = 2.5}$$

v) $\text{Var}(x \mid y = 5) = E(x^2 \mid y = 5) - [E(x \mid y = 5)]^2 = 0.5(1^2 + 4^2) - (2.5)^2 = 8.5 - 6.25 = 2.25$

$$\boxed{\text{Var}(x \mid y = 5) = 2.25}$$

vi) The expected value of y^2 is the summation of each possible squared outcome for y multiplied by the probability of that outcome. Hence:

$$\boxed{E(y^2) = (1/3)(3^2 + 5^2 + 6^2) = 70/3}$$

C. Consider the two functions: $w_1 = 3x^2$ and $w_2 = x^{-1}$. Find: $E(w_1 + w_2)$ and $E(w_1 + w_2 \mid y = 3)$.

Answers: The expected value of a function of x [(i.e., $F(x)$] is equal to the value of the function evaluated at each possible realization of x multiplied by the probability of the associated realizations. Moreover, since the expectation of a sum is the sum of the expectations, $E(w_1 + w_2) = E(3x^2) + Ex^{-1} = 3(1/3)(1^2 + 2^2 + 4^2) + (1/3)(1 + 1/2 + 1/4) = 21 + 1.75/3$. Hence:

$$\boxed{E(w_1 + w_2) = 21.5833}$$

When $y = 3$, x can only take on the values 1 or 2. Hence, $E(w_1 + w_2 \mid y = 3) = E(w_1 \mid y = 3) + E(w_2 \mid y = 3) = 3(1/2)(1^2 + 2^2) + (1/2)(1 + 1/2) = 15/2 + 3/4 = 33/4$. Hence:

$$\boxed{E(w_1 + w_2 \mid y = 3) = 33/4}$$

D. How would your answers change if the balls were drawn with replacement?

Answer: The contingency table changes since it is possible for x and z to take on the same values. The new contingency table becomes:

Summations

	$x = 1$	$x = 2$	$x = 4$
$z = 1$	$y = 2$	$y = 3$	$y = 5$
$z = 2$	$y = 3$	$y = 4$	$y = 6$
$z = 4$	$y = 5$	$y = 6$	$y = 8$

Each entry from (x=1, y=2) through (x=4, y=8) has a probability of 1/9.

Part B becomes:
i) $E(x) = 1/3(1 + 2 + 4) = 7/3$.
ii) $E(y) = (1/9)(2 + 3 + 5 + 3 + 4 + 6 + 5 + 6 + 8) = 42/9 = 14/3$
iii) $E(y \mid x{=}1) = (1/3)(2 + 3 + 5) = 10/3$.
iv) $E(x \mid y{=}5) = (1/2)(1 + 4) = 2.5$.
v) $\text{Var}(x \mid y{=}5) = E(x^2 \mid y{=}5) - [E(x \mid y{=}5)]^2 = (1/2)(1 + 4^2) - (2.5)^2 = 17/2 - 6.25 = 2.25$.
vi). $E(y^2) = (1/9)[2^2 + 2(3)^2 + 4^2 + 2(5)^2 + 2(6)^2 + 8^2] = 224/9$.

Note that the answer to Part C is unchanged.

5. The general solution to an n-th order difference equation requires n arbitrary constants. Consider the second-order equation: $y_t = a_0 + 0.75y_{t-1} - 0.125y_{t-2} + \varepsilon_t$.
A. Find the homogeneous and particular solutions. Discuss the shape of the impulse response function.

Answer: The homogeneous equation is $y_t = 0.75y_{t-1} - 0.125y_{t-2}$. Try the challenge solution $y_t = A\alpha^t$ and obtain
$A\alpha^t - 0.75A\alpha^{t-1} + 0.125A\alpha^{t-2} = 0$

so that the characteristic equation is $\alpha^2 - 0.75\alpha + 0.125 = 0$.

The two characteristic roots are 0.5 and 0.25 and A can be any arbitrary constant. Since a linear combination of the two homogeneous solutions is also a solution, the complete form of the homogeneous solution is $y_t = A_1(0.5)^t + A_2(0.25)^t$ where A_1 and A_2 are arbitrary constants.

The particular solution has the form $y_t = b_0 + \Sigma\alpha_i\varepsilon_{t-i}$. The constant b_0 can easily be found as $b_0 = a_0(1 - 0.75 + 0.125) = 8a_0/3$. To find the α_i substitute $y_t = \Sigma\alpha_i\varepsilon_{t-i}$ into $y_t = 0.75y_{t-1} - 0.125y_{t-2} + \varepsilon_t$ to obtain:

$$\alpha_0\varepsilon_t + \alpha_1\varepsilon_{t-1} + \alpha_2\varepsilon_{t-2} + \alpha_3\varepsilon_{t-3} + ... = 0.75(\alpha_0\varepsilon_{t-1} + \alpha_1\varepsilon_{t-2} + \alpha_2\varepsilon_{t-3} + ...)$$
$$- 0.125(\alpha_0\varepsilon_{t-2} + \alpha_1\varepsilon_{t-3} + \alpha_2\varepsilon_{t-4} + ...) + \varepsilon_t$$

Matching coefficients on like terms, it follows that:

$$\boxed{\alpha_0 = 1;\ \alpha_1 = 0.75;\ \text{and all subsequent } \alpha_i \text{ are such that } \alpha_i = 0.75\alpha_{i-1} - 0.125\alpha_{i-2}}$$

For example, $\alpha_2 = 0.4375$, $\alpha_3 = 0.2348$, $\alpha_4 = 0.1211$, and $\alpha_5 = 0.0615$.

The impulse responses are given by the coefficients of the particular solution. For example, $\partial y_t/\partial \varepsilon_t = 1$; $\partial y_{t+1}/\partial \varepsilon_t = \partial y_t/\partial \varepsilon_{t-1} = 0.75$; $\partial y_{t+2}/\partial \varepsilon_t = \partial y_t/\partial \varepsilon_{t-2} = 0.4375$. Since both characteristic roots are positive and less than unity, the impulse responses converge directly toward the long-run value $y_t = 8a_0/3$.

B. Find the values of the initial conditions (i.e., y_0 and y_1) that ensure $\{y_t\}$ sequence is stationary (Note: A_1 and A_2 are the arbitrary constants in the homogeneous solution).

Answer: The general solution is the sum of the homogeneous and particular solutions:

$$y_t = A_1(0.5)^t + A_2(0.25)^t + 8a_0/3 + \varepsilon_t + 0.75\varepsilon_{t-1} + 0.4375\varepsilon_{t-2} + \ldots$$

so that for periods 0 and 1:

$$y_0 = A_1 + A_2 + 8a_0/3 + \varepsilon_0 + 0.75\varepsilon_{-1} + 0.4375\varepsilon_{-2} + \ldots$$
$$y_1 = A_1(0.5) + A_2(0.25) + 8a_0/3 + \varepsilon_1 + 0.75\varepsilon_0 + 0.4375\varepsilon_{-1} + \ldots$$

The issue is to select y_0 and y_1 such that $A_1 = A_2 = 0$. If $y_0 = 8a_0/3 + \varepsilon_0 + 0.75\varepsilon_{-1} + 0.4375\varepsilon_{-2} + \ldots$ and $y_1 = 8a_0/3 + \varepsilon_1 + 0.75\varepsilon_0 + 0.4375\varepsilon_{-1} + \ldots$, then $A_1 = A_2 = 0$, so that the general solution is:

$$\boxed{y_t = 8a_0/3 + \varepsilon_t + 0.75\varepsilon_{t-1} + 0.4375\varepsilon_{t-2} + \ldots}$$

C. Given your answer to part B, derive the correlogram for the $\{y_t\}$ sequence.

Answer: Given the initial conditions found in Part A, the Yule-Walker equations can be used to derive the correlogram. For simplicity, abstract from the constant a_0 since it does not affect the autocorrelations. Hence, set $a_1 = 0.75$ and $a_2 = -0.125$, and use (2.28) and (2.29) to find the autocorrelations. Given that $\rho_0 = 1$, it follows that $\rho_1 = 0.75 - 0.125\rho_1$ and

$$\boxed{\rho_1 = 2/3 \text{ and } \rho_s = 0.75\rho_{s-1} - 0.125\rho_{s-2}}$$

For example, $\rho_2 = 0.375$; $\rho_3 = 0.1979$; $\rho_4 = 0.1012$; $\rho_5 = 0.0512$.....

6. Consider the second-order stochastic difference equation: $y_t = 1.5y_{t-1} - 0.5y_{t-2} + \varepsilon_t$.

A. Find the characteristic roots of the homogeneous equation.

Answer: The homogeneous equation is $y_t - 1.5y_{t-1} + 0.5y_{t-2} = 0$. If you try the challenge solution $y_t = A\alpha^t$, A and α must satisfy: $A\alpha^t - 1.5A\alpha^{t-1} + 0.5A\alpha^{t-2} = 0$. Dividing by $A\alpha^{t-2}$, the characteristic equation is $\alpha^2 - 1.5\alpha + 0.5 = 0$. Thus, A can be any arbitrary constant and the characteristic roots, (i.e., α) can be 1 or 0.5. The homogeneous solution is:

$$\boxed{y_t = A_1 + A_2(0.5)^t}$$

B. Demonstrate that the roots of $1 - 1.5L + 0.5L^2$ are the reciprocals of your answer in Part A.

Answer: To solve the inverse characteristic equation for L, form $(1 - L)(1 - 0.5L) = 0$. The solutions are $L = 1$ and $L = 1/0.5 = 2$. Hence, the two inverse characteristic roots are the reciprocals of characteristic roots found in Part A. Thus, the stability condition is for the characteristic roots to lie inside of the unit circle or for the roots of the inverse characteristic equation to lie outside of the unit circle.

C. Given initial conditions for y_0 and y_1, find the solution for y_t in terms of the current and past values of the $\{\varepsilon_t\}$ sequence. Explain why it is not possible to obtain the backward looking solution for y_t unless such initial conditions are given.

Answer: One way to solve the problem is to use the initial conditions y_0 and y_1 and iterate forward. Since, $y_2 = 1.5y_1 - 0.5y_0 + \varepsilon_2$, and $y_3 = 1.5y_2 - 0.5y_1 + \varepsilon_3$, it follows that:

$$y_3 = \varepsilon_3 + 1.5\varepsilon_2 + 1.75y_1 - 0.75y_0.$$

Similarly, $y_4 = 1.5y_3 - 0.5y_2 + \varepsilon_4$. Substituting for y_3 and y_2 yields:

$$y_4 = \varepsilon_4 + 1.5\varepsilon_3 + 1.75\varepsilon_2 + 1.875y_1 - 0.875y_0$$

Continuing in this fashion yields:

$$y_5 = \varepsilon_5 + 1.5\varepsilon_4 + 1.75\varepsilon_3 + 1.875\varepsilon_2 + 1.9375y_1 - 0.9375y_0$$
$$y_6 = \varepsilon_6 + 1.5\varepsilon_5 + 1.75\varepsilon_4 + 1.875\varepsilon_3 + 1.9375\varepsilon_2 + 1.96875y_1 - 0.96875y_0$$

The solution has the form:

$$\boxed{y_t = \sum_{i=0}^{t-2} \alpha_i \varepsilon_{t-i} + \alpha_{t-1} y_1 + \alpha_t y_0}$$

where: $\alpha_0 = 1$, $\alpha_1 = 1.5$, $\alpha_t = 1- \alpha_{t-1}$, and the remaining coefficients α_i solve the difference equation $\alpha_i = 1.5\alpha_{i-1} - 0.5\alpha_{i-2}$. Since the coefficients will grow progressively larger, the backward particular solution will not be convergent. The initial conditions are necessary for y_t to be finite for finite values of t.

D. Find the forecast function for y_{t+s}.

Answer: First write the answer to Part C in terms of s rather than in terms of t:

$$y_s = \sum_{i=0}^{s-2} \alpha_i \varepsilon_{s-i} + \alpha_{s-1} y_1 + \alpha_s y_0$$

Then update by t periods to obtain:

$$\boxed{y_{t+s} = \sum_{i=0}^{s-2} \alpha_i \varepsilon_{t+s-i} + \alpha_{s-1} y_{t+1} + \alpha_s y_t}$$

For $s > 1$, forecasts conditioned on the information in $t+1$ and t can be made using $E_{t+1}y_{t+s} = \alpha_{s-1}y_{t+1} + \alpha_s y_t.$

E. Find: Ey_t, Ey_{t+1}, Var(y_t), Var(y_{t+1}), and Cov(y_{t+1}, y_t).

Answers: The point is to illustrate that the $\{y_t\}$ sequence is not stationary. Consider:

i. $Ey_t = \alpha_{t-1}y_1 + \alpha_t y_0$. Since α_{t-1} and α_t are functions of time, the mean is not constant.

ii. $Ey_{t+1} = \alpha_t y_1 + \alpha_{t+1}y_0$. Note that $Ey_{t+1} \neq Ey_t$.

iii. var(y_t) = $[1 + (\alpha_1)^2 + (\alpha_2)^2 + ... + (\alpha_{t-2})^2]\sigma^2$

iv. var(y_{t+1}) = $[1 + (\alpha_1)^2 + (\alpha_2)^2 + ... + (\alpha_{t-1})^2]\sigma^2$ so that var(y_{t+1}) $\neq$ var(y_t).

v. cov(y_{t+1}, y_t) = $[\alpha_0\alpha_1 + \alpha_1\alpha_2 + ... + \alpha_{t-3}\alpha_{t-2}]\sigma^2$.

7. The file entitled SIM_2.XLS contains the simulated data sets used in this chapter. The first column contains the 100 values of the simulated AR(1) process used in Section 7. This first series is entitled Y1. The following programs will perform the tasks indicated in the text. Due to differences in data handling and rounding, your answers need only approximate those presented in the text.

Sample Program for RATS Users

```
all 100                                      ;* The first 3 lines read in the data set
open data a:\sim_2.xls                       ;*  Modify this if your data is not on drive a:
data(format=xls,org=obs)

cor(partial=pacf,qstats,number=24,span=8) y1      ;* calculates the ACF, PACF, and Q-
statistics
graph 1                                           ;* plots the simulated series
# y1

boxjenk(ar=1) y1 / resids                    ;* estimates an AR(1) model and saves the
                                             ;*       residuals in the series called resids

* The next 3 lines compute and display the AIC and SBC
compute aic = %nobs*log(%rss) + 2*%nreg
compute sbc = %nobs*log(%rss) + %nreg*log(%nobs)
display 'aic = ' AIC 'sbc = ' sbc

* Obtain the ACF, PACF, and Q-statistics of the residuals
cor(partial=pacf,qstats,number=24,span=8,dfc=%nreg) resids
```

```
* Now estimate the model with a MA term at lag 12
boxjenk(ar=1,ma=||12||) y1 / resids
cor(partial=pacf,qstats,number=24,span=8,dfc=%nreg) resids
compute aic = %nobs*log(%rss) + 2*%nreg
compute sbc = %nobs*log(%rss) + %nreg*log(%nobs)
display 'aic = ' AIC 'sbc = ' sbc


boxjenk(ar=2) y1 / resids                    ;* Estimates the AR(2) model
boxjenk(ar=1,ma=1) y1 / resids               ;* Estimates an ARMA(1, 1) model
```

Sample Program for STATA Users

```
/*Note: All the code is written for STATA release 8.0. Since STATA cannot read datasets in
the Excel format, please save the Excel datasets as text files and read them in using the line
below, or you can use a program such as Stat/Transfer to translate the datasets from Excel's
format to STATA's format */

clear
cd "X:\New_data"  /*Please change this line accordingly to wherever the dataset is saved*/

insheet using sim_2.txt

tsset obs    /*declare this is a time series dataset*/

*A:plot the sequence against time
line y1 obs

*B:verify the first 12 coefficients of the ACF and PACF
corrgram y1, lag(24)

*C:estimate Model 1
reg y1 l.y1, noconstant

/*the calculation of AIC and SBC makes use of the saved results after estimation commands
such as -reg-, -arima-. type -ereturn list- to obtain a complete list of saved results after a
particular estimation command*/

display "AIC = "=e(N)*ln(e(rss))+2*e(df_m)   /*compute AIC of Model 1*/
display "SBC = "=e(N)*ln(e(rss))+e(df_m)*ln(e(N)) /*compute SBC of Model 1*/
predict resids, residuals  /*calculate residuals after estimation and save it as a new variable called resids*/
corrgram resids, lag(24)  /*verify Q stats of residuals of Model 1*/
```

Page 30: Stationary Models

```
/*Next, reestimate Model 1 using ARIMA command. the reason for doing so is to make the
two models comparable because Model 2 cannot be estimated by linear regression*/

arima y1,ar(1) noconstant
display "AIC = "=-2*e(ll)+2*e(k)
display "SBC = "=-2*e(ll)+e(k)*ln(e(N))
predict resids1,residuals
corrgram resids1, lag(24)  /*verify Q stats of residuals of Model 1*/

/*Note: here the required matsize should be max(AR, MA+1)^2, which is 169(max(1,13)^2).
But matsize for intercooled stata is initially set at 40, far below the required size. Refer to
[R]matsize for details.*/

set matsize 800
arima y1,ar(1) ma(12) noconstant  /*estimate Model 2*/
display "AIC = "=-2*e(ll)+2*e(k)
display "SBC = "=-2*e(ll)+e(k)*ln(e(N))

/*Note: Since ARIMA in STATA estimates a model using a combination of BHHH and
BFGS, it does not report the residual sum of squares which makes impossible the
computation of AIC and SBC using the formula appropriate for a model estimated by OLS.
However, AIC and SBC are still obtainable if we make use of the log likelihood function
value. The final conclusion of which model is better remains unchanged*/

predict resids2, residuals
corrgram resids2, lag(24)  /*verify Q stats of residuals of Model 2*/

*D: estimate the series y1 as an AR(2) without an intercept
reg y1 l.y1 l2.y1, noconstant

*E:estimate the series y1 as an ARMA(1,1) without an intercept
arima y1,ar(1) ma(1) noconstant
```

8. The second column in file entitled SIM_2. XLS contains the 100 values of the simulated ARMA(1, 1) process used in Section 7. This series is entitled Y2. The following programs will perform the tasks indicated in the text. Due to differences in data handling and rounding, your answers need only approximate those reported in the text.

Sample Program for RATS Users

```
all 100                          ;* The first 3 lines read in the data set
open data a:\sim_2.xls           ;*  Modify this if your data is not on drive a:
data(format=xls,org=obs)
```

```
cor(partial=pacf,qstats,number=24,span=8) y2    ;* calculates the ACF, PACF, and Q-statistics
graph 1                                         ;* plots the simulated series
# y2

*RATS contains a procedure to plot autocorrelation and partial autocorrelations.  To use the
*procedure use the following two program lines:

source(noecho) c:\winrats\bjident.src ;* assuming RATS is in a directory called C:\WINRATS
@bjident y2

boxjenk(ar=1) y2 / resids                ;* estimates an AR(1) model and saves the residuals
cor(number=24,partial=partial,qstats,span=8) resids / cors
compute aic = %nobs*log(%rss) + 2*%nreg
compute sbc = %nobs*log(%rss) + %nreg*log(%nobs)
display 'aic = ' AIC 'sbc = ' sbc

boxjenk(ar=2) y2 / resids                ;* estimates an AR(2) model and saves the residuals
cor(number=24,partial=partial,qstats,span=8) resids / cors
compute aic = %nobs*log(%rss) + 2*%nreg
compute sbc = %nobs*log(%rss) + %nreg*log(%nobs)
display 'aic = ' AIC 'sbc = ' sbc

*Continue to copy the statements used in Question 7. For the MA(2) use:

boxjenk(ma=2) y2 / resids
cor(number=24,partial=partial,qstats,span=8) resids / cors

*To compare the MA(2) to the ARMA(1, 1) you need to be a bit careful. For a head-to-head
*comparison, you need to estimate the models over the same sample period. The ARMA(1, 1)
*uses 99 observations while the MA(2) uses all 100 observations. Use:

boxjenk(ma=2) y2 2 100 resids  ;* This instructs RATS to use only observations 2 – 100.
boxjenk(ar=1,ma=1) y2 / resids
cor(number=24,partial=partial,qstats,span=8) resids / cors
compute aic = %nobs*log(%rss) + 2*%nreg
compute sbc = %nobs*log(%rss) + %nreg*log(%nobs)
display 'aic = ' AIC 'sbc = ' sbc
```

Sample Program for STATA Users

```
clear
```

```
cd "a:\New_data" /*Change if the dataset is saved elsewhere*/
insheet using sim_2.txt
tsset obs    /*declare this is a time series dataset*/

*A:plot the sequence against time
line y2 obs
/*verify the first 12 coefficients of the ACF and PACF*/
corrgram y2, lag(12)

*B:estimate Model 1
reg y2 l.y2, noconstant
/*compute AIC and SBC of Model 1(99 obs)*/
display "AIC = "=e(N)*ln(e(rss))+2*e(df_m)
display "SBC = "=e(N)*ln(e(rss))+e(df_m)*ln(e(N))
/*compute AIC and SBC of Model 1(98 obs)*/
display "AIC*= "=(e(N)-1)*ln(e(rss))+2*e(df_m)
display "SBC*= "=(e(N)-1)*ln(e(rss))+e(df_m)*ln(e(N)-1)

predict resids1, residuals
corrgram resids1, lag(24)  /*Q stats of residuals of Model 1*/

arima y2,ar(1) ma(1) noconstant /*estimate Model 2*/
display "AIC = "=-2*e(ll)+2*e(k)
display "SBC = "=-2*e(ll)+e(k)*ln(e(N))

predict resids2, residuals
corrgram resids2, lag(24)  /*verify Q stats of residuals of Model 2*/

reg y2 l.y2 l2.y2, noconstant  /*estimate Model 3*/
display "AIC = "=e(N)*ln(e(rss))+2*e(df_m)
display "SBC = "=e(N)*ln(e(rss))+e(df_m)*ln(e(N))
predict resids3, residuals
corrgram resids3, lag(24)

*C: estimate the process using a pure MA(2) model
arima y2,ma(1 2) noconstant
predict residsc, residuals
corrgram residsc, lag(24)
```

9. The third column in SIM_2.XLS contains the 100 values of a AR(2) process; this series is entitled Y3. The following programs will perform the tasks indicated in the text. Due to differences in data handling and rounding, your answers need only approximate those reported in the text.

Page 33: Stationary Models

Sample Program for RATS Users

Use the first three lines from Question 7 or 8 to read in the data set. To graph the series use:

```
cor(partial=pacf,qstats,number=24,span=8) y3    ;* calculates the ACF, PACF, and Q-statistics
graph 1 ; # y3                                  ;* plots the simulated series

boxjenk(ar=1) y3 / resids                 ;* estimates an AR(1) model and saves the residuals

boxjenk(ar=1,ma=1) y3 / resids        ;* estimates an ARMA(1) model and saves the residuals
cor(partial=pacf,qstats,number=24,span=8,dfc=%nreg) resids

boxjenk(ar=2) y3 / resids                ;* estimates an AR(2) model and saves the residuals
cor(partial=pacf,qstats,number=24,span=8,dfc=%nreg) resids

*To estimate the AR(2) model with the single MA coefficient at lag 16 use:
boxjenk(ar=2,ma=||16||) y3 / resids
```

Sample Program for STATA Users

Use the first four lines from Question 7 or 8 to read in the data set.

To graph the data use:

```
*A:plot the sequence against time
line y3 obs
corrgram y3, lag(12)   /*verify the first 12 coefficients of the ACF and PACF*/

*B:estimate the series as an AR(1) process
reg y3 l.y3, noconstant
display "AIC = "=e(N)*ln(e(rss))+2*e(df_m)   /*compute AIC of the AR(1) model*/
display "SBC = "=e(N)*ln(e(rss))+e(df_m)*ln(e(N)) /*compute SBC of AR(1) model 1*/
predict residsb, residuals
corrgram residsb, lag(24)  /*verify Q stats of residuals of the AR(1) Model*/

*C:estimate the series as an ARMA(1,1) process
arima y3,ar(1) ma(1) noconstant
predict residsc, residuals
corrgram residsc, lag(24)  /*verify Q stats of residuals of the ARMA(1,1) model*/

*E:estimate the series as an AR(2) process
```

```
reg y3 l.y3 l2.y3, noconstant
display "AIC = "=e(N)*ln(e(rss))+2*e(df_m)   /*compute AIC of the AR(2) model*/
display "SBC = "=e(N)*ln(e(rss))+e(df_m)*ln(e(N)) /*compute SBC of the AR(2) model*/
predict residse, residuals
corrgram residse, lag(24)  /*verify Q stats of residuals of the AR(2) model*/

set matsize 800
arima y3,ar(1 2) ma(16) noconstant
```

10. The text uses AIC = T ln(SSR) + 2 n . Suppose that the AIC from model (AIC_1) is smaller than that for model 2 (AIC_2). Hence, model 1 is selected over model 2. In fact, any monotonic transformation of the AIC will always lead to model 1 being selected over model 2. The issue is to show whether the various ways of reporting the AIC (and SBC) are monotonic transformations of each other. Consider the version reported by E-Views:

AIC = $-2\ln(L)/T + 2n/T$ where (as you can see from Appendix 2.1)

$\ln L = -(T/2)\ln(2\pi) - (T/2)\ln\sigma^2 - (2\sigma^2)^{-1}\Sigma(\varepsilon_t)^2$

Since $-(T/2)\ln(2\pi)$ is a constant, it can be ignored in calculating the formula (i.e., adding or subtracting is a monotonic transformation) since it will not affect the relative ranking of the two models. Also note that σ^2 is calculated as $\Sigma(\varepsilon_t)^2/T$. Note that this implies $(\sigma^2)^{-1}\Sigma(\varepsilon_t)^2 = T$ (i.e, a constant). Hence, a monotonic transformation of ln L is

$-(T/2)\ \ln(\Sigma(\varepsilon_t)^2/T)$

Hence, a monotonic transformation of AIC = $-2\ln(L)/T + 2n/T$ is

$\ln(\Sigma(\varepsilon_t)^2/T) + 2n/T$

or multiplying by T and adding ln(T)—two more monotonic transformations, we get

AIC = T ln(SSR) + 2 n

The argument is identical for the SBC and for the formula AIC = exp($2n/T$) (SSR/T).

11. **RATS users can enter the following code:**

```
cal 1960 1 4                    ;* The 4 lines read in the data set
all 2002:1                      ;* The PPI is quarterly over the 1960:1 – 2002:1 period
open data a:\quarterly.xls
data(format=xls,org=obs)
```

```
dif ppi / dy
log ppi / ly                ;* These 2 lines form the logarithmic change
dif ly / dly

*Figure 2.5 was created using:
spg(hfi=1,vfi=3,hea='Figure 2.5')
	gra(hea='Panel a: The Producer Price Index: (1995 = 100)') 1 ; # ppi
	gra(hea='Panel b: First Difference of the PPI') 1 ; # dy
	gra(hea='Panel c: Logarithmic Difference of the PPI') 1 ; # dly
spg(done)

/*The next three lines estimate an ARMA(1, 1), ARMA(1, ||1,4||) and an AR(2) over the full
sample period. To compare the three using the AIC or SBC, estimate all three over the same
sample period.
*/

boxjenk(define=eq1,constant,ar=1,ma=1,dif=1) ly / resids
boxjenk(define=eq1,constant,ar=1,ma=||1,4||,dif=1) ly / resids
boxjenk(define=eq1,constant,ar=2,dif=1) ly / resids

*To perform the 50 out-of-sample forecasts of the ARMA(1, 1), ARMA(1, ||1,4||) models, we
*need to initialize two series to hold the 50 forecasts:

set arma11 = 0.
set arma114 = 0.

*Now each model will be estimated fifty times using the sample period 2 through period 118 + i
*(1989:3 = 2002:1-50). At the end, i = 50 so that the final period is 2001:4. At each estimation,
*the 1-step-ahead forecast is made. The forecasts from the ARMA(1, 1) will be stored in arma11
*and those from the ARMA(1, ||1,4||) will be in arma114.

do i = 1,50
	boxjenk(define=eq1,constant,ar=1,ma=1,noprint) dly * 2002:1-51+i resids
	forecast 1 1 2002:1-50+i
	# eq1 arma11
	boxjenk(define=eq1,constant,ar=1,ma=||1,4||,noprint) dly * 2002:1-51+i resids
	forecast 1 1 2002:1-50+i
	# eq1 arma114
end do i

*Next create the two forecast errors as the actual value of dly minus the forecast
set err11 1989:4 * = dly - arma11
```

set err114 1989:4 * = dly - arma114

*The summary statistics for the two series are given by:
table 1989:4 * err11
table 1989:4 * err114

*For the Diebold-Mariano test, form (2.62) using the absolute values

set d 1989:4 * = abs(err11) - abs(err114)

*The mean value (i.e., $\bar{d}$) and the variance (i.e., γ_0) can be obtained using:

sta d ; com dbar = %mean ; dis dbar %variance

/*You should find that $\bar{d}$ equals -4.99789e-05 and that the variance of d_t is 6.94794e-06. For all practical purposes that autocorrelations of d_t are zero. Hence, the desired statistic is:
*/

$$DM = -4.99789e\text{-}05/[0.00000694794/49)^{0.5}] = -0.13273.$$

* There is not a significant difference between the two forecasts.

*For the Granger-Newbold test, form x and z using
set x 1989:4 * = err11 + err114
set z 1989:4 * = err11 - err114

*Obtain the correlation coefficient using
cross x z 1989:4 * 0 0 rxz

*The correlation is -0.0767318. This is stored in rxz(1). Hence the Granger-Newbold statistic can be computed using:

com den = (1-rxz(1)**2)/49
dis rxz(1)/(den**.5)

*The value of -0.538 is not significant at any conventional level.

Sample Program for STATA Users

clear
cd "a:\New_data" /*Please change this line accordingly if the dataset is saved elsewhere*/
insheet using quarterly.txt
gen obs = _n /*obs is 1,2,...,_N, a new column vector created to act as a time variable*/

```
tsset obs    /*declare this is a time series dataset*/
gen lppi=ln(ppi)
gen dlppi=d.lppi  /*make the data transformation*/

*A:estimate the series as an ARMA(1,1) process
arima dlppi, ar(1) ma(1)
predict resids1, residuals
corrgram resids1, lag(12)  /*verify Q stats of residuals of the ARMA(1,1) Model*/

arima dlppi, ar(1) ma(1 4) /*estimate the series as an ARMA(1,1 4) process*/
predict resids2, residuals
corrgram resids2, lag(12)  /*verify Q stats of residuals of the ARMA(1,1 4) Model*/

reg dlppi l.dlppi l2.dlppi  /*estimate the series as an AR(2) process*/
display "AIC = "=e(N)*ln(e(rss))+2*e(df_m)   /*compute AIC of the AR(2) model*/
display "SBC = "=e(N)*ln(e(rss))+e(df_m)*ln(e(N)) /*compute SBC of AR(2) model 1*/
predict resids3, residuals
corrgram resids3, lag(12)  /*verify Q stats of residuals of the AR(2) Model*/

*B:estimate the ARMA(1,1) model over the period 1960:Q1 to 1989:Q3
gen time = real(substr(descriptor,1,4) + substr(descriptor,6,1)) /*this line is to change the string
variable "descriptor" in the original dataset to a numerical variable "time" to facilitate the
following estimation over certain periods*/

arima dlppi if time <= 19893, ar(1) ma(1)
predict e1t if time > 19893, residuals /*obtain the fifty one-step ahead forecast errors from the
ARMA(1,1) model and save it to the new variable e1t*/
summ e1t

*B:estimate the ARMA(1,1 4) model over the period 1960:Q1 to 1989:Q3
arima dlppi if time <= 19893, ar(1) ma(1 4)
predict e2t if time > 19893, residuals /*obtain the fifty one-step ahead forecast errors from the
ARMA(1,1 4) model*/
summ e2t

*C:DM test using mean absolute error to compare the ARMA(1,1) to the ARMA(1,(1,4)) model
gen yt=abs(e1t)-abs(e2t) if time > 19893
summ yt if time > 19893,detail
display "DM Stat="=r(mean)/sqrt(r(Var)/(r(N)-1))

*D:Granger-Newbold test to compare the AR(2) model to the ARMA(1,1) model
reg dlppi l.dlppi l2.dlppi if time <= 19893 /*first obtain the fifty one-step ahead forecast errors
from the AR(2) model*/
```

```
predict e3t if time > 19893, residuals

gen xt=e1t+e3t if time > 19893
gen zt=e1t-e3t if time > 19893
cor xt zt
display "GN Stat="=r(rho)/sqrt((1-(r(rho))^2)/49)

*D:Granger-Newbold test to compare the ARMA(1,1) model to the ARMA(1,(1,4)) model
gen xt1=e1t+e2t if time > 19893
gen zt1=e1t-e2t if time > 19893
cor xt1 zt1
display "GN Stat="=r(rho)/sqrt((1-(r(rho))^2)/49)
```

12. **RATS Users can use the following to obtain the results reported in Section 11.**

```
*The first 4 lines read in the data set
cal 1960 1 4
all 2002:1
open data a:\quarterly.xls
data(format=xls,org=obs)
```

*Next create *x* as M1NSA, and make a number of transformations

```
set x = m1nsa
log x / lx                      ;*Log transformation = lx
dif lx / dlx                    ;* Difference of the log = dlx
dif(sdiffs=1) lx / slx          ; * First seasonal difference of the log = slx
dif(sdiffs=1,dif=1) lx / sdlx   ;* First seasonal difference and regular difference of the log = sdlx
set growth = 100*dlx            ;* Equivalent to 100*log(x/x{1})

* Figures 2.7 was created with
label growth x                          ;* created the legend for the graph
# 'Rate of Growth' 'M1 in Billions'

gra(hea='Figure 2.7: The Level and Growth Rate of M1',overlay=line, $
        patterns,key=below,nokbox,ovl='billions of $',vla='rate of growth') 2 ;# growth  ; # x;

*The AR(1) model with a seasonal MA(1) coefficient was estimated using:
box(ar=1,sma=1,constant) sdlx / resids

*The AR(1) model with a multiplicative seasonal AR(1) coefficient was estimated using:
box(ar=1,sar=1,constant) sdlx / resids
```

*The MA(1) model with a multiplicative seasonal MA(1) coefficient was estimated using:
box(ma=1,sma=1,constant) sdlx / resids

B. Now use DDNSA. Simply redefine *x* as *ddnsa* and make the indicated transformations:

```
set x = ddnsa
log x / lx
dif lx / dlx
dif(sdiffs=1) lx / slx
dif(sdiffs=1,dif=1) lx / sdlx
set growth = dlx                    ;* Equivalent to log(x/x{1})
```

*If you form the correlation of *x*, you should obtain the indicated values. The ACF and PACF of the growth rate (i.e., dlx) is given by

```
cor(number=24,partial=partial,qstats,span=4) dlx / cors
```

C. The correlations of the log($DDNSA_t$) – log($DDNSA_{t-4}$) is given by
cor(number=12,partial=partial,qstats,span=4) slx

D. Estimate the three models for sdlx over the same sample period. The first model has the lowest AIC and SBC. The residuals from this model show no sign of serial correlation.

```
box(ar=1,sma=1,constant) sdlx 1962:3 * resids
cor(number=24,partial=partial,qstats,span=4) resids / cors
box(ar=1,sar=1,constant) sdlx 1962:3 * resids
cor(number=24,partial=partial,qstats,span=4) resids / cors
box(ma=1,sma=1,constant) sdlx 1962:3 * resids
cor(number=24,partial=partial,qstats,span=4) resids / cors
```

CHAPTER 3

MODELING VOLATILITY

Lecture Suggestions

1. I make an overhead transparency of Figure 3-7 to illustrate ARCH processes. As described on page 116 of the text, the upper-left-hand graph shows 100 serially uncorrelated and normally distributed disturbances representing the $\{v_t\}$ sequence. These disturbances were used to construct the $\{\varepsilon_t\}$ sequence shown in the upper-right-hand graph. Each value of ε_t was constructed using the formula $\varepsilon_t = v_t[1 + 0.8(\varepsilon_{t-1})^2]$. The lower two graphs show the interaction of the ARCH error term and the magnitude of the AR(1) coefficients. Increasing the magnitude of the AR(1) coefficient from 0.0, to 0.2, to 0.9, increased the volatility of the simulated $\{y_t\}$ sequence. For your convenience, a copy of the figure is reproduced below.

2. Instead of assigning Question 4 as a homework assignment, I use the three models to illustrate the properties of an ARCH-M process. Consider the following three models:

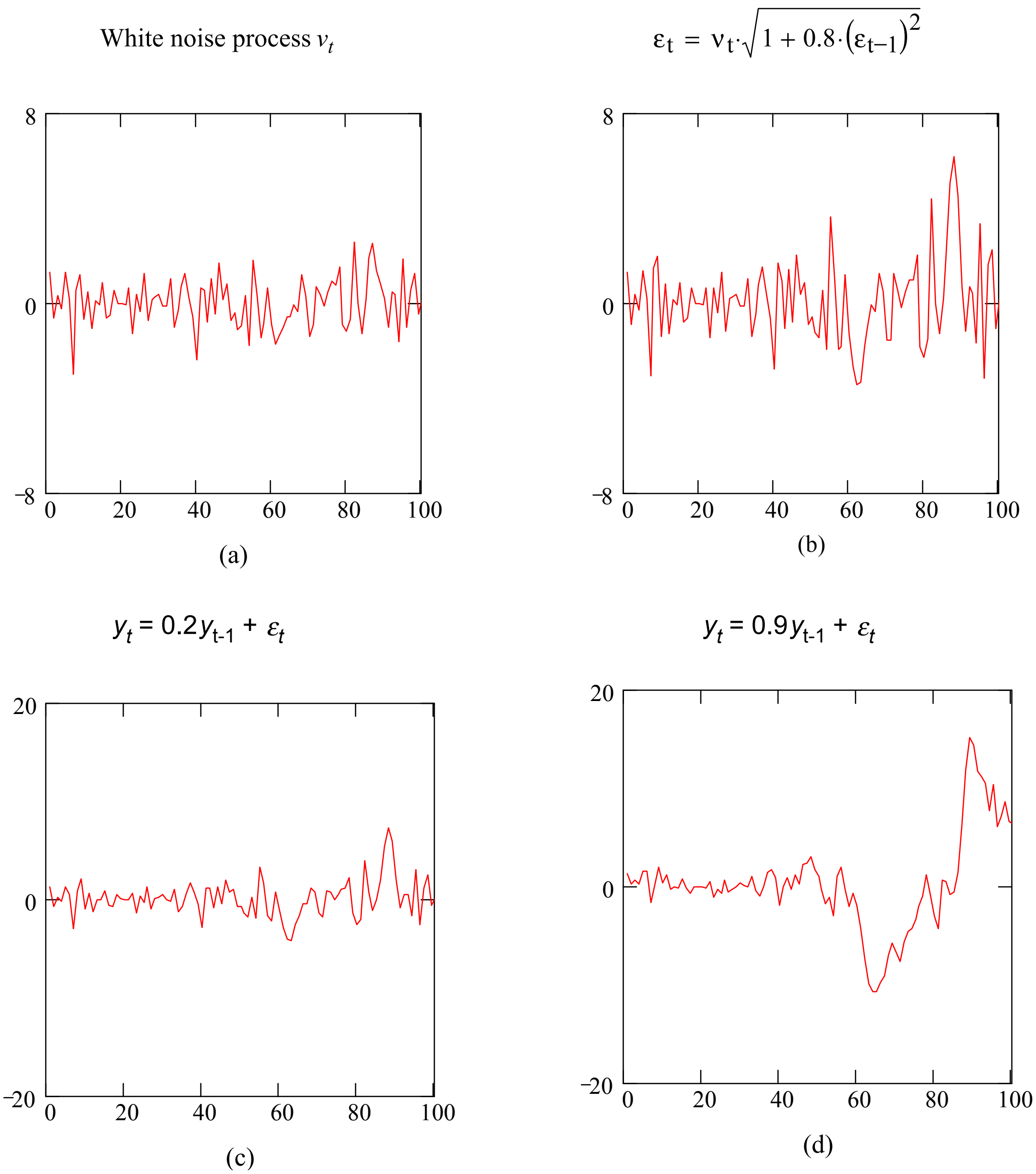
Model 1: $y_t = 0.5y_{t-1} + \varepsilon_t$

Model 2: $y_t = \varepsilon_t - (\varepsilon_{t-1})^2$

Model 3: $y_t = 0.5y_{t-1} + \varepsilon_t - (\varepsilon_{t-1})^2$

Model 1 is a pure AR(1) process that is familiar to the students. Model 2 is a pure ARCH-M process. When the realized value of ε_{t-1} is large in absolute value, the conditional expectation of y_t is negative: $E_{t-1}y_t = -(\varepsilon_{t-1})^2$. Thus, Model 2 illustrates a simple process in which the conditional mean is

Figure 3.7: Simulated ARCH Processes



negatively related to the absolute value of the previous period's error term. Suppose that all values of ε_i for $i \le 0$ are zero. Now, if the next 5 values of the ε_t sequence are (1, -1, -2, 1, 1), y_t has the time path shown in Figure 3M-1 (see the answer to Question 4 below). I use an overhead transparency of Figure 3M-1 to compare the path of the AR(1) and ARCH-M models. Model 3 combines the AR(1) model with the ARCH-M effect exhibited by Model 2. The dotted line shown in Figure 3M-1 shows how the AR(1) and ARCH-M effects interact.

Answers to Questions

1. Suppose that the $\{\varepsilon_t\}$ sequence is the ARCH(q) process

$$\varepsilon_t = v_t \cdot [\alpha_0 + \alpha_1(\varepsilon_{t-1})^2 + ... + \alpha_q (\varepsilon_{t-q})^2]^{1/2}$$

Show that $E_{t-1}\varepsilon_t^2$ has the same form as the conditional variance of (3.1).

Answer: Regardless of whether or not the actual regression residuals are used, the expected value of $E(\varepsilon_t)^2$ is:

$$E_{t\text{-}1}(\varepsilon_t)^2 = \alpha_0 + \alpha_1 E(\varepsilon_{t\text{-}1})^2 + ... + \alpha_q E(\varepsilon_{t\text{-}q})^2$$

Now using (3.8), $\varepsilon_t = v_t \cdot [\alpha_0 + \alpha_1(\varepsilon_{t-1})^2 + ... + \alpha_q (\varepsilon_{t-q})^2]^{1/2}$ so that:

$$\begin{aligned} E_{t\text{-}1}(\varepsilon_t)^2 &= E_{t\text{-}1}[(v_t)^2 (\alpha_0 + \alpha_1 (\varepsilon_{t\text{-}1})^2 + ... + \alpha_q (\varepsilon_{t\text{-}q})^2)] \\ &= E_{t\text{-}1}(v_t)^2 \cdot E_{t\text{-}1}[\alpha_0 + \alpha_1 (\varepsilon_{t\text{-}1})^2 + ... + \alpha_q (\varepsilon_{t\text{-}q})^2] \\ &= \alpha_0 + \alpha_1 (\varepsilon_{t\text{-}1})^2 + ... + \alpha_q (\varepsilon_{t\text{-}q})^2 \end{aligned}$$

Thus, using either (3.1) or (3.8):

$$\boxed{E_{t\text{-}1}(\varepsilon_t)^2 = \alpha_0 + \alpha_1(\varepsilon_{t\text{-}1})^2 + ... + \alpha_q (\varepsilon_{t\text{-}q})^2}$$

2. Consider the ARCH-M model represented by equations (3.23) through (3.25). Recall that $\{\varepsilon_t\}$ is a white noise disturbance; for simplicity let $E(\varepsilon_t)^2 = E(\varepsilon_{t\text{-}1})^2 = ... = 1$.

A. Find the unconditional mean: Ey_t. How does a change in δ affect the mean? Using the example of section 6, show that changing β and δ from (-4, 4) to (-1, 1) preserves the mean of the $\{y_t\}$ sequence.

Answer: Combine (3.23), (3.24) and (3.25) and take the expectation of the result to obtain:

$$Ey_t = E(\mu_t + \varepsilon_t) = E(\beta + \delta h_t + \varepsilon_t) = E[\beta + \delta[\alpha_0 + \sum_{i=1}^{q} \alpha_i (\varepsilon_{t\text{-}i})^2] + \varepsilon_t]$$

Since $E(\varepsilon_t)^2 = E(\varepsilon_{t\text{-}1})^2 = ... = 1$, and $E\varepsilon_t = 0$ it follows that:

$$\boxed{Ey_t = \beta + \delta(\alpha_0 + \sum_{i=1}^{q} \alpha_i)}$$

Increasing β has a 1:1 effect on the mean; the effect of a one-unit change in δ in Ey_t is α_0 +

$\Sigma\alpha_i$. It is straightforward to show that $\alpha_0 + \Sigma\alpha_i = 1$. Since $E(\varepsilon_t)^2 = E[(v_t)^2 h_t] = 1$ and $E(v_t)^2 = 1$, it follows that $E(h_t) = 1$. Hence: $Eh_t = E[\alpha_0 + \Sigma\alpha_i(\varepsilon_{t-i})^2] = \alpha_0 + \Sigma\alpha_i = 1$. Given $\beta = -4$, $\delta = 4$, the unconditional mean of y_t is:

$$Ey_t = -4 + 4(\alpha_0 + \Sigma\alpha_i) = -4 + 4 \cdot 1 = 0.$$

Similarly, if $\beta = -1$ and $\delta = 1$, then $Ey_t = 0$. Hence, changing β and δ from (-4, 4) to (-1, 1) preserves the mean of the $\{y_t\}$ sequence.

B. Show that the unconditional variance of y_t when $h_t = \alpha_0 + \alpha_1(\varepsilon_{t-1})^2$ does not depend on β, δ, and α_0.

Answer: From Equation (3.23), $y_t = \mu_t + \varepsilon_t$ so that the unconditional variance of y_t is:

$$\text{Var}(y_t) = \text{var}(\mu_t) + \text{var}(\varepsilon_t) + 2\,\text{Cov}(\mu_t, \varepsilon_t).$$

Note that: $\text{Var}(\varepsilon_t) = E(\varepsilon_t)^2 - [E(\varepsilon_t)]^2 = 1 - 0^2 = 1$. Next, form $\text{cov}(\mu_t, \varepsilon_t)$ as:

$$\begin{aligned}
\text{Cov}(\mu_t, \varepsilon_t) &= E(\mu_t\varepsilon_t) - E(\mu_t) \cdot E(\varepsilon_t) \\
&= E[(\beta + \delta \text{h}_t) \cdot \varepsilon_t] - E(\beta + \delta \text{h}_t) \cdot 0 \\
&= E\{[\beta + \delta \cdot (\alpha_0 + \alpha_1(\varepsilon_{t-1}^{2})] \cdot \varepsilon_t\} \\
&= E[\beta\varepsilon_t + \delta \cdot (\alpha_0 + \alpha_1(\varepsilon_{t-1})^2)\,\varepsilon_t] \\
&= \beta E(\varepsilon_t) + \delta \cdot \text{E}[(\alpha_0 + \alpha_1(\varepsilon_{t-1})^2)\varepsilon_t] \\
&= \beta 0 + \delta 0 = 0
\end{aligned}$$

Now, find $\text{Var}(\mu_t)$ as:

$$\begin{aligned}
\text{Var}(\mu_t) &= \text{var}(\beta + \delta h_t) = \delta^2 \text{var}(h_t) = \delta^2 \text{var}[\alpha_0 + \alpha_1(\varepsilon_{t-1})^2] \\
&= \delta^2(\alpha_1)^2 \text{var}[(\varepsilon_{t-1})^2].
\end{aligned}$$

Hence, the unconditional variance of y_t does not depend on β and α_0 since $\text{var}(y_t) = \text{var}(\mu_t + \varepsilon_t)$ and $\text{cov}(u_t, \varepsilon_t) = 0$. Thus:

$$\boxed{\text{Var}(y_t) = 1 + \delta^2(\alpha_1)^2 \cdot \text{var}[(\varepsilon_{t-1})^2]}$$

However, as opposed to the assertion in the question, increasing the absolute value of δ increases $\text{var}(y_t)$.

3. Bollerslev proved that the ACF of the squared residuals resulting from the GARCH (p, q) process represented by (3.9) acts as an ARMA (m, p) process where $m = \max(p, q)$. You are to illustrate this result using the examples below.

A. Consider the GARCH(1,2) process: $h_t = \alpha_0 + \alpha_1(\varepsilon_{t-1})^2 + \alpha_2 (\varepsilon_{t-2})^2 + \beta_1 h_{t-1}$. Following the steps in the question, it is possible to write:

$$(\varepsilon_t)^2 = \alpha_0 + (\alpha_1 + \beta_1)(\varepsilon_{t-1})^2 + \alpha_2 (\varepsilon_{t-2})^2 - \beta_1 \eta_{t-1} + \eta_t$$
where: $\eta_t = (\varepsilon_t)^2 - h_t$

Show that η_t is serially uncorrelated and that the $\{(\varepsilon_t)^2\}$ sequence acts as an ARMA(2,1) process.

Answer: Recall that: $\varepsilon_t = v_t \sqrt{h_t}$ where $E(v_t) = 0$ and $\text{Var}(v_t) = 1$. To show that η_t is serially uncorrelated, we need to prove $\text{Cov}(\eta_t, \eta_s) = 0$, for $t \neq s$. By definition, $\eta_t = (\varepsilon_t)^2 - h_t$, and $\varepsilon_t = v_t \sqrt{h_t}$ so that:

$$\eta_t = (\varepsilon_t)^2 - h_t = (v_t)^2 h_t - h_t = [(v_t)^2 - 1]\, h_t.$$

Since v_t and h_t are independent, it follows that:

$$E(\eta_t) = E\{[(v_t)^2 - 1]\, h_t\} = E\,[(v_t)^2 - 1]E(h_t) = (1-1) \cdot E(h_t) = 0, \forall t.$$

$$\begin{aligned} \text{As such, } \text{cov}(\eta_t, \eta_s) &= E\,\{[\eta_t - E(\eta_t)][\eta_s - E(\eta_s)]\} \\ &= E\,(\eta_t \cdot \eta_s) = E\,\{[(v_t)^2-1]h_t \cdot [(v_s)^2-1]h_s\} \end{aligned}$$

$$\boxed{\text{cov}(\eta_t, \eta_s) = E[(v_t)^2-1]E(h_t) \cdot E[(v_s)^2-1]E(h_s) = 0}$$

Thus, η_t is serially uncorrelated. To show that the $\{(\varepsilon_t)^2\}$ sequence acts as an ARMA(2,1) process use the results that:

$E(\eta_t) = 0$ and
$\text{Cov}(\eta_t, \eta_s) = 0$, for $t \neq s$.

Since $(\varepsilon_t)^2 = \alpha_0 + (\alpha_1+\beta_1)(\varepsilon_{t-1})^2 + \alpha_2 (\varepsilon_{t-2})^2 - \beta_1\eta_{t-1} + \eta_t$, it is immediately clear that the $\{(\varepsilon_t)^2\}$ sequence acts as an ARMA(2,1) process with autoregressive coefficients $(\alpha_1+\beta_1)$, α_2, a moving average coefficient $-\beta_1$, and $\{\eta_t\}$ as a white noise process.

B. Consider the GARCH(2, 1) process: $h_t = \alpha_0 + \alpha_1(\varepsilon_{t-1})^2 + \beta_1 h_{t-1} + \beta_2 h_{t-2}$. Show that it is possible to add η_t to each side to obtain:

$$(\varepsilon_t)^2 = \alpha_0 + \alpha_1(\varepsilon_{t-1})^2 + \beta_1 h_{t-1} + \eta_t + \beta_2 h_{t-2}.$$

Show that adding and subtracting the terms $\beta_1(\varepsilon_{t-1})^2$ and $\beta_2 (\varepsilon_{t-2})^2$ to the right-hand-side of this equation yields an ARMA(2,2) process.

Answer: Given the GARCH(2, 1) process, $h_t = \alpha_0 + \alpha_1(\varepsilon_{t-1})^2 + \beta_1 h_{t-1} + \beta_2 h_{t-2}$, it is possible to add $\eta_t = (\varepsilon_t)^2 - h_t$ to each side to obtain:

$$h_t + [(\varepsilon_t)^2 - h_t] = \alpha_0 + \alpha_1(\varepsilon_{t-1})^2 + \beta_1 h_{t-1} + \beta_2 h_{t-2} + (\varepsilon_t)^2 - h_t.$$

Hence: $(\varepsilon_t)^2 = \alpha_0 + \alpha_1 (\varepsilon_{t-1})^2 + \eta_t + \beta_1 h_{t-1} + \beta_2 h_{t-2}$. Now, add and subtract the terms $\beta_1(\varepsilon_{t-1})^2$ and $\beta_2(\varepsilon_{t-2})^2$ to the right-hand-side of above equation yields:

$$\begin{aligned}(\varepsilon_t)^2 &= \alpha_0 + \alpha_1(\varepsilon_{t-1})^2 + \eta_t + \beta_1 h_{t-1} + \beta_2 h_{t-2}\\ &= \alpha_0 + \alpha_1(\varepsilon_{t-1})^2 + \eta_t + \beta_1 h_{t-1} - \beta_1(\varepsilon_{t-1})^2 + \beta_1(\varepsilon_{t-1})^2 + \beta_2 h_{t-2} - \beta_2(\varepsilon_{t-2})^2 + \beta_2(\varepsilon_{t-2})^2\\ &= \alpha_0 + \alpha_1(\varepsilon_{t-1})^2 + \eta_t - \beta_1[(\varepsilon_{t-1})^2 - h_{t-1}] + \beta_1(\varepsilon_{t-1})^2 - \beta_2 [(\varepsilon_{t-2})^2 - h_{t-2}] + \beta_2(\varepsilon_{t-2})^2\end{aligned}$$

$$\boxed{(\varepsilon_t)^2 = \alpha_0 + (\alpha_1 + \beta_1)(\varepsilon_{t-1})^2 + \beta_2(\varepsilon_{t-2})^2 + \eta_t - \beta_1\eta_{t-1} - \beta_2\eta_{t-2}}$$

Since $\{\eta_t\}$ acts as a white noise process, the $\{(\varepsilon_t)^2\}$ sequence acts as an ARMA(2,2) process with AR coefficients $(\alpha_1 + \beta_1)$ and β_2 and MA coefficients $(-\beta_1)$, and $(-\beta_2)$.

C. Provide an intuitive explanation of the statement:
"The Lagrange Multiplier for ARCH errors test cannot be used to test the null of white noise squared residuals against an alternative of a specific GARCH(p,q) process."

Answer: Every GARCH(p,q) process has a representation of ARMA(m,p) process for the ACF of the squared residuals, where m=max(p,q). So that the Lagrange Multiplier for ARCH errors test cannot be used to test the null of white noise squared residuals against an alternative of a specific GARCH(p,q) process.

D. Sketch the proof of the general statement that the ACF of the squared residuals resulting from the GARCH(p,q) process represented by (3.9) acts as an ARMA(m,p) process where m=max(p,q).

Answer: Consider the GARCH(p,q) process: $h_t = \alpha_0 + \alpha_i \sum_{i=1}^{q} (\varepsilon_{t-i})^2 + \sum_{j=1}^{p} \beta_j \mathrm{h}_{t-j}$.

Add the expression $(\varepsilon^2_t - h_t)$ to each side to give:

$$\begin{aligned}(\varepsilon_t)^2 &= \alpha_0 + \sum_{i=1}^{q} \alpha_i(\varepsilon_{t-i})^2 + \sum_{j=1}^{p} \beta_j h_{t-j} + (\varepsilon_t)^2 - h_t\\ &= \alpha_0 + \sum_{i=1}^{q} \alpha_i(\varepsilon_{t-i})^2 + [\sum_{j=1}^{p} \beta_j(\varepsilon_{t-j})^2 - \sum_{j=1}^{p} \beta_j(\varepsilon_{t-j})^2] + \sum_{j=1}^{p} \beta_j h_{t-j} + (\varepsilon_t)^2 - h_t\\ &= \alpha_0 + [\sum_{i=1}^{q} \alpha_i (\varepsilon_{t-i})^2 + \sum_{j=1}^{p} \beta_j(\varepsilon_{t-j})^2] - \sum_{j=1}^{p} \beta_j[(\varepsilon_{t-j})^2 - h_{t-j}] + (\varepsilon_t)^2 - h_t\\ &= \alpha_0 + (\alpha + \beta_1)(\varepsilon_{t-1})^2 + (\alpha_2 + \beta_2) (\varepsilon_{t-2})^2 + \ldots + (\alpha_m + \beta_m)(\varepsilon_{t-m})^2\\ &\quad -\beta_1\eta_{t-1} - \beta_2\eta_{t-2} - \ldots - \beta_p \eta_{t-p} + \eta_t,\end{aligned}$$

where $\eta_t \equiv (\varepsilon_t)^2$ - h_t, m = max(p,q), $\alpha_i \equiv 0$ for $i > q$, and $\beta_j \equiv 0$ for $j > p$. Hence:

$$\boxed{\begin{gathered}(\varepsilon_t)^2 = \alpha_0 + \sum_{k=1}^{m}(\alpha_k + \beta_k)(\varepsilon_{t-k})^2 + \eta_t - \beta_1\eta_{t-1} - \beta_2\eta_{t-1} - \beta_2\eta_{t-2} - ... - \beta_p\,\eta_{t-p},\\ \textit{where } m = \max(p,q),\ \alpha_i \equiv 0 \text{ for } i > q,\ \beta_j \equiv 0 \text{ for } j > p.\end{gathered}}$$

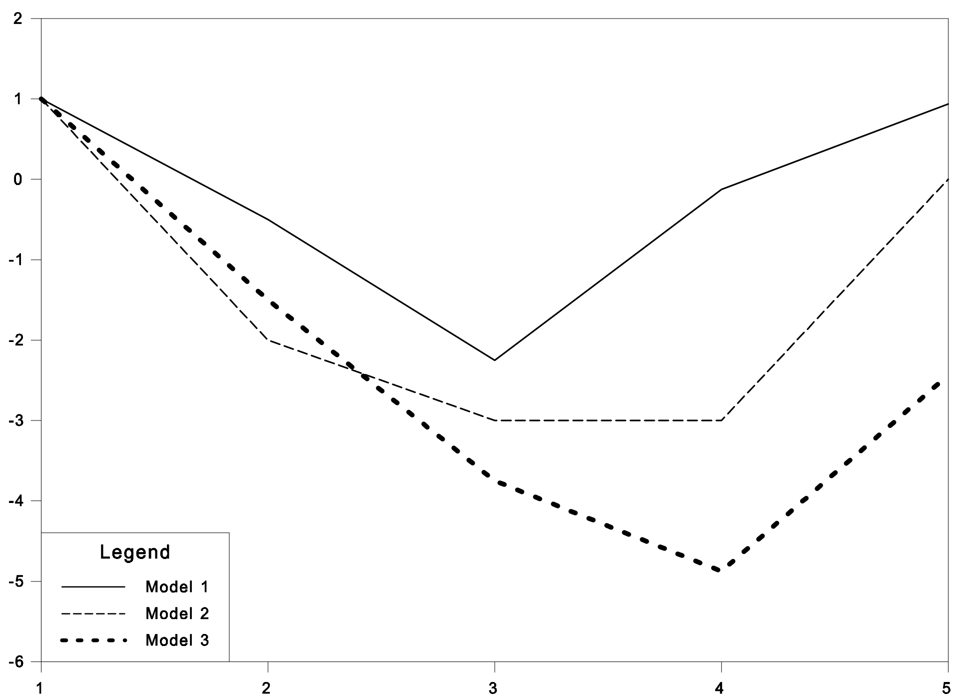
Since η_t is serially uncorrelated the ACF of the squared residuals resulting from the GARCH(p, q) process acts as an ARMA(m, p) process where m = max(p, q).

4. Let $y_0 = 0$ and let the first five realizations of the $\{\varepsilon_t\}$ sequence be: (1,-1,-2, 1, 1). Plot each of the following sequences:

Model 1: $y_t = 0.5y_{t-1}+\varepsilon_t$
Model 2: $y_t = \varepsilon_t - (\varepsilon_{t-1})^2$
Model 3: $y_t = 0.5y_{t-1} + \varepsilon_t - (\varepsilon_{t-1})^2$

Answer: The time paths of the three models are shown in Figure 3M-1. Note that Model 3 contains autoregressive and ARCH-M effects. The effect of introducing the autoregressive term in going from Model 2 to Model 3 increases the volatility of the series. Similarly, the effect of introducing the ARCH-M effect in going from Model 1 to Model 3 is to increase the volatility of the series. The autoregressive coefficients and ARCH-M effects interact.

Figure 3M-1: An ARCH-M Process



B. For each of the three models, calculate the sample mean and variance of $\{y_t\}$.

Answer: For Model 1, the realizations are: (1,-0.5,-2.25,-0.125, 0.9375). The sample mean of Model 1 is: (1/5) (1 - 0.5 - 2.25 -0.125 + 0.9375) = -0.1875, and the sample variance is: (1/5) $[1^2 +(-0.5)^2+(-2.25)^2 +(-0.125)^2+(0.9375)^2] - (-0.1875)^2 = 1.40625$.

For Model 2, the realizations are: (1,-2,-3,-3, 0). The sample mean of Model 2 is: (1/5) (1-2-3-3+0) = -1.4, and the sample variance is (1/5) $[1^2 +(-2)^2+(-3)^2+(-3)^2+0] - (-1.4)^2 = 2.64$.

For Model 3, the realizations are: (1,-1.5,-3.75,-4.875, -2.4375). The sample mean is (1/5) (1-1.5-3.75-4.875-2.4375) = -2.3125 and the sample variance is: (1/5) $[1^2+ (-1.5)^2 + (-3.75)^2 + (-4.875)^2 + (-2.4375)^2] - (-2.3125)^2 = 4.05625$.

Hence:

Model	Mean	Variance
1	-0.1875	1.40625
2	-1.4	2.64
3	-2.3125	4.05625

5. The file labeled ARCH.XLS contains the 100 realizations of the simulated $\{y_t\}$ sequence used to create the lower right-hand panel of Figure 3.10. The following program will reproduce the reported results.

Sample Program for RATS Users:

```
all 100                        ;* allocates space for 100 observations
open data a:\arch.xls          ;* opens the data set assumed to be on drive a:\
data(format=xls,org=obs)

table / y                ;* produces the summary statistics for the ARCH(1) series
                         ;* ARCH(1) series labeled y

* Next, estimate an AR(1) model without an intercept and produce the ACF and PACF.
boxjenk(ar=1) y / resids
cor(partial=pacf,qstats,number=24,span=4,dfc=1) resids

*  Now, define sqresid as the squared residuals from the AR(1) model and construct the ACF
* and PACF of these squared residuals.
set sqresid = resids*resids
cor(partial=pacf,qstats,number=24,span=4,dfc=1) sqresid
```

```
linreg sqresid                  ;* estimate an AR(1) model of the squared residuals
# constant sqresid{1}

compute trsq= %nobs*%rsquared   ;* Calculate TR^2 and its significance level
cdf chisqr trsq 1

linreg sqresid                  ;* estimate an AR(4) model of the squared residuals.
# constant sqresid{1 to 4}

nonlin b1 a0 a1                 ;*  prepares for a non-linear estimation b1 a0 and a1
frml regresid = y - b1*y{1}     ;*  defines the residual
frml archvar = a0 + a1*regresid(t-1)**2       ;* defines the variance
frml archlogl = (v=archvar(t)), -0.5*(log(v)+regresid(t)**2/v)  ;* defines the likelihood
boxjenk(ar=1) y                 ;* estimate an AR(1) in order to obtain an initial
compute b1=%beta(1)             ;*      guess for the value of b1 and a0
compute a0=%seesq, a1=.3        ;* the initial guesses of a0 and a1

*  Given the initial guesses and the definition of archlogl, the next line performs the non-
*  linear estimation of b1, a0, and a1.
maximize(method=bhhh,recursive,iterations=75) archlogl 3 *
```

Sample Programs for STATA Users

```
clear
cd "a:\New_data" /*Please change this line accordingly if the dataset is saved elsewhere*/
insheet using arch.txt
summarize y
gen obs = _n /*obs is 1,2,...,_N, a new column vector created to act as a time variable*/
tsset obs  /*declare this is a time-series dataset*/

/*A:estimate the series using OLS*/
reg y l.y,noconstant
predict resids,residuals
corrgram resids,lag(24) /*B: obtain the ACF and PACF of the residuals*/

gen sqresid=resids^2 /*C:obtain the ACF and PACF of the squared residuals*/
corrgram sqresid,lag(24)

/*D:estimate the squared residuals as an AR(1) process*/
reg sqresid l.sqresid
scalar define lm = e(N)*e(r2)  /*make use of the saved results after estimation to calculate test statistics*/
display "Lagrange Multiplier is: "=lm
```

```
display "Prob-value is: "=chi2tail(e(df_m),lm)

/*E:estimate the squared residuals as an ARCH(4) process*/
reg sqresid l.sqresid l2.sqresid l3.sqresid l4.sqresid

/*F:simultaneously estimate the sequence and the ARCH(1) error process using ML estimation*/
arch y l.y, arch(1) noconstant
```

6. The file QUARTERLY.XLS contains the quarterly values of the U.S. Producer Price Index (PPI) that were used in Section 4. The following program will estimate the various GARCH models of the logarithmic change in the PPI.

```
cal 1960 1 4                    ;* Sets the calendar dates to begin in 1960:Q1
all 2002:1                      ;* through 2002:Q1
open data a:\quarterly.xls      ;* Change this line if the data is not on drive a:
data(format=xls,org=obs)

* The next three lines form the inflation rate as the logarithmic change in the PPI
dif ppi / dy
log ppi / ly
dif ly / dly

*The next line estimate the ARMA(1, ||1,4||) model and stores the residuals in the series resids. The
* following lines creates the correlations and Q-statistics of resids.
boxjenk(define=eq1,constant,ar=1,ma=||1 , 4 ||) dly / resids
cor(qstats,number=24,span=4) resids

* Next, create the squared residuals. The correlations of the squared residuals and the Q-statistics are
* obtained using
set r2 = resids**2
cor(qstats,span=4,number=12) r2

* Estimate a linear regression in the form of (3.17) using four lags of the squared residuals
lin r2 ; # constant r2{ 1 to 4}

* Now use eight lags
lin(nopri) r2 ; # constant r2{1 to 8}

* To exclude lags 5 through 8 use:
exc ; # r2{5 to 8}

* The following block of instructions can be used to estimate the ARMA(1,(1,4)) model with
* ARCH(4) errors. Additional details can be found in Chapter 1 of the Programming Manual.
```

```
set u = 0.0
nonlin b0 b1 b2 b3 a0 a1 a2 a3 a4
frml e = dly - b0 - b1*dly{1} - b2*u{1} - b3*u{4}
frml var = a0 + a1*e{1}**2 + a2*e{2}**2 + a3*e{3}**2 + a4*e{4}**2
frml L = (u = e), -.5*(log(var)+e(t)**2/var)
boxjenk(noprint,constant,ar=1,ma=||1,4||) dly
compute b0=%beta(1), b1=%beta(2), b2 = %beta(3) , b3 = %beta(4)
compute a0=%seesq, a1=.1 , a2 = .1 , a3 = .6 , a4 = .1
maximize(iterations=75) L 6 *


* To constrain the coefficients in the ARCH(4) process to be 0.4, 0.3, 0.2 and 0.1, use
set u = 0.0
nonlin b0 b1 b2 b3 a0 a1
frml e = dly - b0 - b1*dly{1} - b2*u{1} - b3*u{4}
frml var = a0 + a1*(.4*e{1}**2 + .3*e{2}**2 + .2*e{3}**2 + .1*e{4}**2)
frml L = (u = e), -.5*(log(var)+e(t)**2/var)
boxjenk(noprint,constant,ar=1,ma=||1,4||) dly
compute b0=%beta(1), b1=%beta(2), b2 = %beta(3) , b3 = %beta(4)
compute a0=%seesq, a1=.1 , a2 = .1 , a3 = .6 , a4 = .1
maximize(iterations=75) L 6 *

* To estimate the ARMA(1,(1,4)) with GARCH(1,1) errors, use
set w = 0.0
set u = 0.0
nonlin b0 b1 b2 b3 a0 a1 a2
frml e = dly - b0 - b1*dly{1} - b2*u{1} - b3*u{4}
frml var = a0 + a1*e{1}**2 + a2*w{1}
frml L = (u = e), (w = var), -.5*(log(var)+e(t)**2/var)
boxjenk(noprint,constant,ar=1,ma=||1,4||) dly
compute b0=%beta(1), b1=%beta(2), b2 = %beta(3) , b3 = %beta(4)
compute a0=%seesq, a1=.3 , a2 = .5
maximize(iterations=75) L 6 *
```

*The residuals are contained in the series *u*. You can perform the appropriate diagnostics on this series.

STATA Users can use the following program

```
clear
cd "a:\New_data" /*Please change this line accordingly if the dataset is saved elsewhere*/
insheet using quarterly.txt
gen obs = _n /*obs is 1,2,...,_N, a new column vector created to act as a time variable*/
tsset obs
```

```
gen piet=ln(ppi)-ln(l.ppi)

/*A:estimate equation (3.16)*/
arima piet,ar(1) ma(1 4)
predict resids,residuals
corrgram resids,lag(24) /*obtain the ACF and PACF of the residuals*/

/*B:plot the ACF and PACF of the squared residuals*/
gen r2 = resids^2
corrgram r2,lag(12)
ac r2,lag(12)
pac r2,lag(12)

reg r2 l.r2
reg r2 l.r2 l2.r2 l3.r2 l4.r2 /*estimate a linear regression in the form of (3.17) using four lags of
the squared residuals*/

test l2.r2 l3.r2 l4.r2  /*test if an ARCH(1) model is adequate for characterizing the error
process*/
quietly reg r2 l.r2 l2.r2 l3.r2 l4.r2 l5.r2 l6.r2 l7.r2 l8.r2  /*use eight lags*/
test l5.r2 l6.r2 l7.r2 l8.r2  /*test if an ARCH(4) model is necessary to characterize the error
process*/

/*C:replicate the results for the various GARCH models in section 4*/
/*due to the different methods used in STATA to maximize the likelihood function and the
possible difference in the initial values of the coefficients, the following regression results are
slightly different from those in RATS*/
arch piet,ar(1) ma(1 4) arch(1/4) /*estimate piet using an ARMA(1,(1,4)) model assuming an
ARCH(4) error process*/

/*constrain the coefficients in the ARCH(4) process to be 0.4,0.3,0.2,0.1*/
constrain define 1 (3/4)*[ARCH]l1.arch=[ARCH]l2.arch
constrain define 2 (2/4)*[ARCH]l1.arch=[ARCH]l3.arch
constrain define 3 (1/4)*[ARCH]l1.arch=[ARCH]l4.arch
arch piet,ar(1) ma(1 4) arch(1/4) constraint(1/3)
lincom [ARCH]l.arch/.4  /*recover the alpha parameter from the original specification. Any of
the four arch parameters could be used to produce an identical estimate. It is equivalent to A1 in
RATS.*/

arch piet,ar(1) ma(1 4) arch(1) garch(1) /*estimate piet using an ARMA(1,(1,4)) model assuming
an GARCH(1,1) error*/
```

7. The second series on the file **ARCH.XLS** contains 100 observations of a simulated ARCH-M

process. The following programs will produce the indicated results.

```
* The next three lines read in the 100 observations from the file.
all 100
open data a:/arch.xls
data(format=xls,org=obs)

table / y_m                ;* The second series on the file is called y_m. TABLE produces the desired
                           ;* summary statistics.

* The following instruction produces the graph of the ARCH-M process
graph(header='Simulated ARCH-M Process') 1 ; # y_m

* To estimate the MA(||3,6||) process and save the residuals as resids, use:
boxjenk(constant,ma=||3,6||) y_m / resids

*The correlations of resids (and the Q-statistics) are given by
cor(partial=pacf,qstats,number=24,span=8) resids

* Now form the squared residuals and obtain the autocorrelations using
set ressq = resids*resids                         ;* Form the squared residuals
cor(partial=pacf,qstats,number=24,span=4) ressq

* Perform the Lagrange multiplier test for ARCH(4) errors by regressing the squared error on its own
* four lags.
linreg ressq; # constant ressq{1 to 4}

*To obtain the F-statistic use
exclude ; # ressq{1 to 4}

* Alternatively, to obtain TR^2 use
compute trsq = %nobs*%rsquared     ;* Compute T*R2 and obtain the cumulative density
cdf chisqr trsq 4                  ;* of trsq as chi-square with 4 degrees of freedom.

set u = 0.0
nonlin b0 b1 a0 a1
frml var = a0 + a1*u{1}**2                     ;* variance equation
frml e = y_m - b0 - b1*var(t)                  ;* mean equation
frml L = (u = e), -.5*(log(var)+e**2/var)      ;* likelihood function
compute b0=.8, b1=.6  , a0=.2, a1=.7           ;* initial guesses
maximize L 2 *
```

Sample Program for STATA Users

```
clear
```

```
cd "X:\New_data" /*Please change this line accordingly if the dataset is saved elsewhere*/
insheet using arch.txt
summarize y_m
gen obs = _n /*obs is 1,2,..., _N, a column vector that acts as a time variable*/
tsset obs /*declare this is a time series dataset*/

line y_m obs  /*produce the graph of the ARCH-M process*/
/*estimate the MA(3,6) process and store the residuals in the new variable resids*/
set matsize 800 /*matsize is initially set at 40 but here it requires at least max(ar,ma+1)^2 or
49(max(0,6+1)^2). see [R]matsize for more detail*/
arima y_m,ma(3 6)
predict resids,residuals
corrgram resids,lag(24)
gen ressq = resids^2
corrgram ressq,lag(24)

/*perform the Lagrange Multiplier test for ARCH(4) errors*/
reg ressq l.ressq l2.ressq l3.ressq l4.ressq
test l.ressq l2.ressq l3.ressq l4.ressq  /*obtain the F-stat of the null hypothesis of no ARCH
errors*/
display "LM stat ="=e(N)*e(r2)  /*make use of the saved results after estimation to calculate test
statistics*/
display "Prob-value is: "=chi2tail(e(df_m),e(N)*e(r2))

/*C,D:estimate y series as the ARCH-M process*/
arch y_m,archm arch(1)
predict resids1,residuals
corrgram resids1,lag(12)
```

8. Consider the ARCH(2) process: $E_{t-1}(\varepsilon_t)^2 = \alpha_0 + \alpha_1(\varepsilon_{t-1})^2 + \alpha_2(\varepsilon_{t-2})^2$.

A. Suppose that the residuals come from the model $y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$. Find the conditional and unconditional variance of $\{y_t\}$ in terms of the parameters a_1, α_0, α_1, and α_2.

Answer: To find the conditional variance of $\{y_t\}$, first note that the conditional mean of y_t is: $E_{t-1}y_t = E_{t-1}(a_0+a_1y_{t-1}+\varepsilon_t) = a_0+a_1 \cdot y_{t-1}$. The conditional variance of $\{y_t\}$ is:

$$\begin{aligned}
\mathrm{Var}(y_t|y_{t-1}, y_{t-2},...) &= E_{t-1}[y_t - E_{t-1}(y_t)]^2 \\
&= E_{t-1}\,[(a_0+a_1y_{t-1}+\varepsilon_t) - (a_0+a_1y_t)]^2 = E_{t-1}(\varepsilon_t)^2 \\
&= E_{t-1}[\alpha_0 + \alpha_1(\varepsilon_{t-1})^2 + \alpha_2(\varepsilon_{t-2})^2] \\
&= \alpha_0 + \alpha_1(\varepsilon_{t-1})^2 + \alpha_2(\varepsilon_{t-2})^2
\end{aligned}$$

To find the unconditional variance of $\{y_t\}$ find the particular solution for y_t as:

$y_t = a_0/(1\text{-}a_1) + \varepsilon_t + a_1\varepsilon_{t-1} + (a_1)^2\varepsilon_{t-2} + ...$

Thus, the unconditional mean is: $Ey_t = a_0/(1\text{-}a_1)$. Using the solution for y_t and Ey_t, the unconditional variance of y_t is:

$$\begin{aligned}\text{Var}(y_t) = E\,[y_t\text{-}E(y_t)]^2 &= E[\varepsilon_t + a_1\varepsilon_{t-1} + (a_1)^2\varepsilon_{t-2} + (a_1)^3\varepsilon_{t-3} + ...\,]^2 \\ &= \sigma^2/[1 - (a_1)^2]\end{aligned}$$

Given that $\varepsilon_t = v_t \sqrt{\alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2}$, it follows that $(\varepsilon_t)^2 = v_t^2[\alpha_0+\alpha_1(\varepsilon_{t-1})^2 + \alpha_2(\varepsilon_{t-2})^2)$. Since $E\varepsilon_t = Ev_t = 0$:

$$\begin{aligned}E(\varepsilon_t)^2 &= E\,(v_t)^2E[\alpha_0 + \alpha_1(\varepsilon_{t-1})^2+\alpha_2(\varepsilon_{t-2})^2] \\ &= \alpha_0+\alpha_1\,E(\varepsilon_{t-1})^2 + \alpha_2\,E(\varepsilon_{t-2})^2\end{aligned}$$

Since the unconditional variance of ε_t is identical to that of ε_{t-1} and ε_{t-2}, the unconditional variance is:

$$E(\varepsilon_t)^2 = \alpha_0\,/(1\text{-}\alpha_1\text{-}\alpha_2)$$

so that the unconditional variance of y_t is:

$$\boxed{\text{Var}(y_t) = \sum_{i=0}^{\infty}(a_1)^{2i}\text{Var}(\varepsilon_{t-i}) = \alpha_0\,/[(1\text{-}\alpha_1\text{-}\alpha_2)(1\text{-}(a_1)^2)].}$$

B. Suppose that $\{y_t\}$ is an ARCH-M process such that the level of y_t is positively related to its own conditional variance. For simplicity, let: $y_t = \alpha_0 + \alpha_1(\varepsilon_{t-1})^2 + \alpha_2(\varepsilon_{t-2})^2 + \varepsilon_t$. Trace out the impulse response function of $\{y_t\}$ to an ε_t shock. You may assume that the system has been in long-run equilibrium ($\varepsilon_{t-2} = \varepsilon_{t-1} = 0$) but now $\varepsilon_1 = 1$. Thus, the issue is to find the values of y_1, y_2, y_3, and y_4 given that $\varepsilon_2 = \varepsilon_3 =...= 0$.

Answer: Iterating forward from the initial conditions:

$$\begin{aligned}y_1 &= \alpha_0 + \alpha_1(\varepsilon_0)^2 + \alpha_2(\varepsilon_{-1})^2 + \varepsilon_1 = \alpha_0 + \alpha_1\cdot 0^2 + \alpha_2\cdot 0^2 + 1 = \alpha_0+1 \\ y_2 &= \alpha_0 + \alpha_1(\varepsilon_1)^2 + \alpha_2(\varepsilon_0)^2 + \varepsilon_2 = \alpha_0 + \alpha_1 + \alpha_2\cdot 0^2 + 0 = \alpha_0+\alpha_1 \\ y_3 &= \alpha_0 + \alpha_1(\varepsilon_2)^2 + \alpha_2(\varepsilon_1)^2 + \varepsilon_3 = \alpha_0 + \alpha_1\cdot 0^2 + \alpha_2\cdot 1^2 + 0 = \alpha_0+\alpha_2 \\ y_4 &= \alpha_0 + \alpha_1(\varepsilon_3)^2 + \alpha_2(\varepsilon_2)^2 + \varepsilon_4 = \alpha_0 + \alpha_1\cdot 0^2 + \alpha_2\cdot 0^2 + 0 = \alpha_0 \\ y_5 &= \alpha_0 + \alpha_1(\varepsilon_4)^2 + \alpha_2(\varepsilon_3)^2 + \varepsilon_5 = \alpha_0 + \alpha_1\cdot 0^2 + \alpha_2\cdot 0^2 + 0 = \alpha_0,....\end{aligned}$$

Hence:

$$\boxed{dy_1/d\varepsilon_1 = 1;\ dy_2/d\varepsilon_1 = \alpha_1;\ dy_3/d\varepsilon_1 = \alpha_2;\ \text{and } dy_i/d\varepsilon_1 = 0 \text{ for } i > 3.}$$

C. Use your answer to Part B to explain the following result. A student estimated $\{y_t\}$ as an MA(2) process and found the residuals to be serially uncorrelated. A second student estimated the same series as the ARCH-M process: $y_t = \alpha_0 + \alpha_1(\varepsilon_{t-1})^2 + \alpha_2(\varepsilon_{t-2})^2 + \varepsilon_t$. Why might both estimates appear reasonable? How would you decide which is the better model?

> **Answer**: If α_1 and α_2 are positive, positive values of ε_t will be associated with positive values of y_t, y_{t+1}, and y_{t+2}. Negative realizations of ε_t will be associated with positive values of y_{t+1} and y_{t+2}. In any event, y_{t+1} and y_{t+2} will always appear to move together. However, there is no particular tendency for y_{t+1} and y_{t+3} to move together. The sequence may appear to be a MA process. Diagnostic checking of the squares of the residuals from the MA(2) model will indicate if there are ARCH-M effects. The point (see question 8*d*) is that an ARCH-M process may appear to be an MA process

9. This program produces the results for the NYSE data used in Section 10.

Sample Program for RATS Users

Instead of programming the maximum likelihood estimation, this section illustrates the GARCH.SRC procedure. You can download the procedure from www.estima.com.

```
* To begin, the first four lines read in the data set NYSE.XLS. Note that the data set begins on
January 3, 1995 and ends on August 30, 2002. The typical week has five observations. However,
the market closes for events such as holidays and September 11, 2001. No attempt has been made
to adjust the dates for such missing observations. As such, the data set contains 1938
observations; hence it appears to end on May 28, 2001.

CAL(daily) 1995 1 3
all 2002:9:30
open data a:\nyse.xls
data(org=obs,format=xls)

*Create the summary statistics with
tab

* Create the logarithmic change in the NYSE Composite Index using the next two instructions.
* The resulting variables is called rate.
set x = composite; set rate = log(x)-log(x{1})

cor(qstats,number=24,span=4) rate          ;* Obtain the ACF and Q-statistics of rate

* To obtain the two equations in (3.41) use
box(constant,ma=1) rate                    ;* Estimate the MA(1)
box(constant,ma=2) rate / resids           ;* Estimate the MA(2) and save the residuals
```

```
cor(qstats,number=24,span=4) resids          ;* Obtain the ACF of the residuals from the MA(2)
set r2 = resids**2                           ;* Create the squared residuals called r2.
cor(qstats,number=24,span=4) r2              ;* Obtain the ACF of r2.

* The next two lines perform the LM test for ARCH errors be regressing the squares residual on
* its own four lags. The F-statistic is produced by the EXCLUDE instruction
lin r2 ; # constant r2{1 to 4}
exclude ; # r2{1 to 4}

* GARCH.SRC will estimate most GARCH-type models. Compile the procedure using
source(noecho) c:\winrats\garch.src

* Respectively, estimate MA(2)-GARCH(1,1) and the MA(1)-GARCH(1,1) models with
@garch(mod=garch,ma=2,p=1,q=1,nointeractive) rate
@garch(mod=garch,ma=1,p=1,q=1,nointeractive) rate

* Estimate the IGARCH and ARCH-M models with
@garch(mod=igarch,ma=1,p=1,q=1,nogra,nointeractive) rate
@garch(mod=garch,ma=1,p=1,q=1,nointeractive,aim,term=none,iters=200) rate

* Estimate the IGARCH model have the naormalized residuals in nres and the conditional
* variance in cvar (as described on page 147, h_t = cvar and nres = ε_t/(h_t)^0.5
@garch(mod=igarch,ma=1,p=1,q=1,nointeractive,nogra) rate / nres cvar

cor(qstats,number=24,span=4) nres   ;* Check for remaining serial correlation in nres
set n2 = nres*nres                  ;* Square the normalized residuals
lin n2 ; # constant n2{1 to 4}      ;* LM test for remaining ARCH effects in the residuals

* As on page 147, check for leverage effects by regressing n2 on the lagged normalized residuals.
lin n2 ; # constant nres{1 to 3}

* To perform the Engle-Ng test, create the indicator function
set d1 = %if(nres<0,1.,0.)      ;* d1 is 1 if the standardized residual is negative
lin n2 ; # constant d1{1}       ;* regress n2 on the lagged value of the indicator

* For the positive and negative sign bias test multiple d1 and (1-d1) by the standardized residuals
set xminus = d1*nres ; set xplus = (1-d1)*nres
lin n2 ;#  constant d1{1} xplus{1} xminus{1}          ;* estimate the regression

* Estimate the EGARCH model
@garch(mod=egarch,ma=1,p=1,q=1,nointeractive) rate / nres cvar
```

Sample Program for STATA Users

```
clear
```

```
cd "X:\New_data" /*Please change this line accordingly if the dataset is saved elsewhere*/
insheet using arch.txt
summarize y_m
gen obs = _n /*obs is 1,2,...,_N, a column vector that acts as a time variable*/
tsset obs /*declare this is a time series dataset*/

line y_m obs  /*produce the graph of the ARCH-M process*/
/*estimate the MA(3,6) process and store the residuals in the new variable resids*/
set matsize 800 /*matsize is initially set at 40 but here it requires at least max(ar,ma+1)^2 or
49(max(0,6+1)^2). see [R]matsize for more detail*/
arima y_m,ma(3 6)
predict resids,residuals
corrgram resids,lag(24)
gen ressq = resids^2
corrgram ressq,lag(24)

/*perform the Lagrange Multiplier test for ARCH(4) errors*/
reg ressq l.ressq l2.ressq l3.ressq l4.ressq
test l.ressq l2.ressq l3.ressq l4.ressq  /*obtain the F-stat of the null hypothesis of no ARCH
errors*/
display "LM stat ="=e(N)*e(r2)  /*make use of the saved results after estimation to calculate test
statistics*/
display "Prob-value is: "=chi2tail(e(df_m),e(N)*e(r2))

/*C,D:estimate y series as the ARCH-M process*/
arch y_m,archm arch(1)
predict resids1,residuals
corrgram resids1,lag(12)
```

10. The each step of the estimation is explained in the Programming Manual.

CHAPTER 4

MODELS WITH TREND

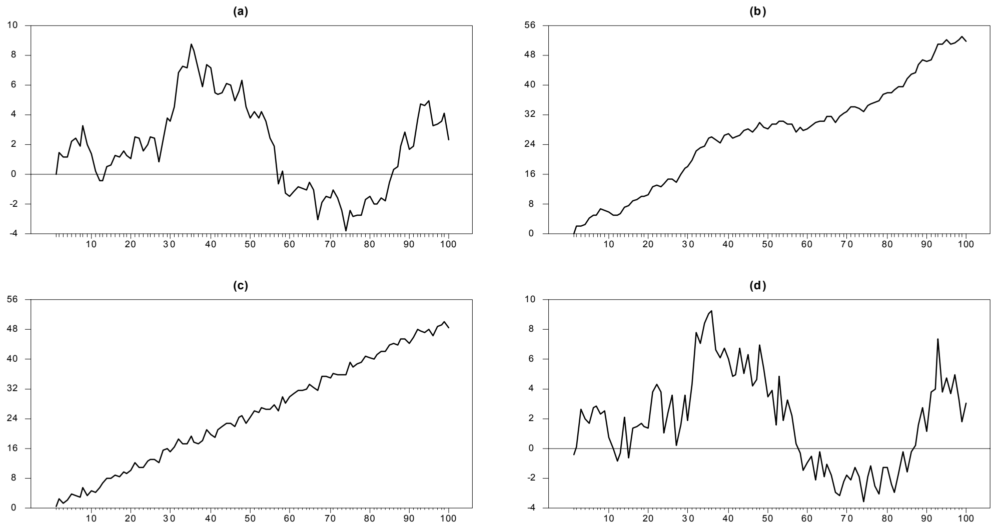
Lecture Suggestions

1. A common misconception is that it is possible to determine whether or not a series is stationary by visually inspecting the time path of the data. I try to dispel this notion using overhead transparencies of the interest rates shown in Figure 4.2 and the four graphs in Figure 4.3. I cover-up the headings in figures and ask the students if they believe that the series are stationary. Most will argue that interest rates cannot behave as nonstationary processes. However, they will not be able to discern any difference between the interest rate series and panels (a) and (d) of Figure 4.3.

All agree that the two series in graph (b) and (c) of Figure 4.3 are non-stationary. However, there is no simple way to determine whether the series are trend stationary or difference stationary. I use these same overheads to explain why unit root tests have very low power. Figure 4.3 (with captions the captions removed) is reproduced on the next page for your convenience.

2. Much of the material in Chapter 4 relies on the material in Chapter 1. I remind students of the relationship between characteristic roots, stability, and stationarity. At this point, I solve some of the mathematical questions involving unit root process. You can select from Question 5, 6, 9 and 10 of Chapter 1 and Question 1 of Chapter 4.

Figure 4.3: Four Series With Trends



Answers to Questions

1. Given an initial condition for y_0, find the solution for y_t. Also find the s-step-ahead forecast $E_t y_{t+s}$

A. $y_t = y_{t-1} + \varepsilon_t + 0.5\ \varepsilon_{t-1}$

Answer: Iterating from y_0 forward:

$y_1 = y_0+\varepsilon_1+0.5\ \varepsilon_0$,
$y_2 = y_1+\varepsilon_2+0.5\ \varepsilon_1 = (y_0+\varepsilon_1+0.5\ \varepsilon_0) + \varepsilon_2+0.5\ \varepsilon_1 = y_0+\varepsilon_2+1.5\ \varepsilon_1+0.5\ \varepsilon_0$,
$y_3 = y_2+\varepsilon_3+0.5\ \varepsilon_2 = (y_0+\varepsilon_2+1.5\ \varepsilon_1+0.5\ \varepsilon_0) + \varepsilon_3+0.5\ \varepsilon_2$
$= y_0+\varepsilon_3+1.5\ \varepsilon_2+1.5\ \varepsilon_1+0.5\ \varepsilon_0$,
$y_4 = y_3 + \varepsilon_4 + 0.5\varepsilon_3 = (y_0 + \varepsilon_3 + 1.5\varepsilon_2 + 1.5\varepsilon_1 + 0.5\varepsilon_0) + \varepsilon_4 + 0.5\varepsilon_3$
$= y_0 + \varepsilon_4 + 1.5\varepsilon_3 + 1.5\varepsilon_2 + 1.5\varepsilon_1+0.5\varepsilon_0$,
…
$y_t = y_0+\varepsilon_t+1.5\ (\varepsilon_{t-1}+\varepsilon_{t-2} +...+ \varepsilon_2+\varepsilon_1) + 0.5\ \varepsilon_0$.

Update the solution for y_t by s periods so that the solution for y_{t+s} can be written as:

$y_{t+s} = y_t + \varepsilon_{t+s} + 1.5(\varepsilon_{t+s-1} + \varepsilon_{t+s-2} +...+ \varepsilon_{t+2} + \varepsilon_{t+1}) + 0.5\ \varepsilon_t$.

The forecast function is: $\boxed{E_t\, y_{t+s} = y_t + 0.5\,\varepsilon_t}$

Hence, the value of $(y_t+0.5\varepsilon_t)$ is the unbiased estimator of all future values of y_{t+s} for all s>o.

B. $y_t = 1.1\, y_{t-1} + \varepsilon_t$

Answer: As in Part A, iterate forward to obtain:

$y_1 = 1.1\, y_0 + \varepsilon_1,$
$y_2 = 1.1\, y_1 + \varepsilon_2 = 1.1\,(1.1\, y_0+\varepsilon_1) + \varepsilon_2 = 1.1^2\, y_0 + \varepsilon_2 + 1.1\,\varepsilon_1,$
$y_3 = 1.1\, y_2 + \varepsilon_3 = 1.1\,(1.1^2\, y_0+\varepsilon_2+1.1\varepsilon_1) + \varepsilon_3 = 1.1^3\, y_0 + \varepsilon_3 + 1.1\,\varepsilon_2 + 1.1^2\,\varepsilon_1,$
$y_4 = 1.1\, y_3 + \varepsilon_4 = 1.1\,(1.1^3\, y_0+\varepsilon_3+1.1\varepsilon_2+1.1^2\varepsilon_1) + \varepsilon_4 = 1.1^4\, y_0 + \varepsilon_4$
$\quad + 1.1\,\varepsilon_3 + 1.1^2\,\varepsilon_2+1.1^3\varepsilon_1,$
…
$y_t = 1.1^t\, y_0 + \varepsilon_t + 1.1\,\varepsilon_{t-1} + 1.1^2\varepsilon_{t-2} +\ldots+ 1.1^{t-2}\varepsilon_2 + 1.1^{t-1}\,\varepsilon_1.$

Update by s periods so that the solution for y_{t+s} can be written as:

$y_{t+s} = 1.1^s y_t + \varepsilon_{t+s} + 1.1\,\varepsilon_{t+s-1} + 1.1^2\varepsilon_{t+s-2} +\ldots+ 1.1^{s-2}\varepsilon_{t+2} + 1.1^{s-1}\,\varepsilon_{t+1}.$

The forecast function is: $E_t y_{t+s} = 1.1^s y_t$

In contrast to part (a), the forecast function is not flat. As s increases, the forecast value will grow at the rate of 1.1. Note also that this process is not stationary since it explodes.

C. $y_t = y_{t-1}+1+\varepsilon_t$

Answer: Iterating forward:

$y_1 = y_0 +1+\varepsilon_1,$
$y_2 = y_1+1+\varepsilon_2 = (y_0 +1+\varepsilon_1) + 1+\varepsilon_2 = y_0 +2+\varepsilon_2+\varepsilon_1,$
$y_3 = y_2+1+\varepsilon_3 = (y_0 +2+\varepsilon_2+\varepsilon_1) +1+\varepsilon_3 = y_0 +3+\varepsilon_3+\varepsilon_2+\varepsilon_1,$
$y_4 = y_3+1+\varepsilon_4 = (y_0 +3+\varepsilon_3+\varepsilon_2+\varepsilon_1) +1+\varepsilon_4 = y_0 +4+\varepsilon_4+\varepsilon_3+\varepsilon_2+\varepsilon_1,$
…

$$y_t = y_0 + t + \sum_{i=1}^{t} \varepsilon_i$$

The solution for y_{t+s} can be written as:

$$y_{t+s} = y_t + s + \sum_{i=1}^{s} \varepsilon_{t+i}.$$

Taking the expectation, the forecast function is:

$E_t y_{t+s} = y_t + s$

The mean change in y_t is always the constant 1, is reflected in the forecast function. This is a random-walk plus drift model with drift term equal to 1.

D. $y_t = y_{t-1}+t+\varepsilon_t$

Answer: Using the same technique as in A through C.

$y_1 = y_0 +1+\varepsilon_1,$
$y_2 = y_1+2+\varepsilon_2 = (y_0 +1+\varepsilon_1) +2+\varepsilon_2 = y_0 +1+2+\varepsilon_2+\varepsilon_1,$
$y_3 = y_2+3+\varepsilon_3 = (y_0 +1+2+\varepsilon_2+\varepsilon_1) +3+\varepsilon_3 = y_0 +1+2+3+\varepsilon_3+\varepsilon_2+\varepsilon_1,$
$y_4 = y_3+4+\varepsilon_4 = (y_0 +1+2+3+\varepsilon_3+\varepsilon_2+\varepsilon_1) +4+\varepsilon_4 = y_0+1+2+3+4+\varepsilon_4+\varepsilon_3+\varepsilon_2+\varepsilon_1,$
…
$y_t = y_0 +(1+2+3 +...+ t) + (\varepsilon_t+\varepsilon_{t-1} +...+ \varepsilon_2+\varepsilon_1).$

The solution for y_{t+s} can be written as:

$y_{t+s} = y_t+[(t+1)+(t+2)+...+(t+s)] + (\varepsilon_{t+s}+\varepsilon_{t+s-1}+...+\varepsilon_{t+2}+\varepsilon_{t+1}).$

The forecast function is:

$$\boxed{E_t\, y_{t+s} = y_t+(t+1)+(t+2)+...+(t+s) = y_t + s\cdot t + \frac{s(s+1)}{2}}$$

Note that the change in y_t is positively related to t. As such, the slope of the forecast function is positively related to t.

E. $y_t = \mu_t + \eta_t + 0.5\eta_{t-1}$, where: $\mu_t = \mu_{t-1}+\varepsilon_t$

Answer: Given initial conditions for all stochastic terms:

$y_1 = \mu_1 + \eta_1 + 0.5\eta_0 = \mu_0 + \varepsilon_1 + \eta_1 + 0.5\eta_0,$
$y_2 = \mu_2+\eta_2+0.5\eta_1 = (\mu_1+\varepsilon_2) +\eta_2+0.5\eta_1 = \mu_0+\varepsilon_1+\varepsilon_2+\eta_2+0.5\eta_1,$
$y_3 = \mu_3+\eta_3+0.5\eta_2 = (\mu_2+\varepsilon_3) +\eta_3+0.5\eta_2 = \mu_0+\varepsilon_1+\varepsilon_2+\varepsilon_3+\eta_3+0.5\eta_2,$
$y_4 = \mu_4+\eta_4+0.5\eta_3 = (\mu_3+\varepsilon_4) +\eta_4+0.5\eta_3 = \mu_0+\varepsilon_1+\varepsilon_2+\varepsilon_3+\varepsilon_4+\eta_4+0.5\eta_3,$
–
$y_t = \mu_0+\varepsilon_1+\varepsilon_2+...+\varepsilon_{t-1}+\varepsilon_t+\eta_t+0.5\eta_{t-1}.$

In period zero, the value of y_0 is given by: $y_0 = \mu_0+\eta_0+0.5\eta_{-1}$ (assume that μ_0, η_0 and η_{-1} are known) so that we can write the model as:

$$y_t = (y_0-\eta_0-0.5\eta_{-1}) + \sum_{i=1}^{t} \varepsilon_i+\eta_t+0.5\eta_{t-1}.$$

The solution for y_{t+s} can be written as:

$$y_{t+s} = (y_t-\eta_t-0.5\eta_{t-1}) + \sum_{i=1}^{s} \varepsilon_{t+i} +\eta_{t+s}+0.5\eta_{t+s-1}.$$

Hence, the forecast function is:

$$\boxed{\begin{aligned} &E_t y_{t+1} = y_t - 0.5\eta_t - 0.5\eta_{t-1}, \quad s=1, \\ &E_t y_{t+s} = y_t - \eta_t - 0.5\eta_{t-1}, \quad s > 1. \end{aligned}}$$

Here y_t is a random-walk plus a stationary component that is a MA(1) process.

F. $y_t = \mu_t + \eta_t + 0.5\eta_{t-1}$, where $\mu_t = 0.5 + \mu_{t-1} + \varepsilon_t$.

Answer:

$y_1 = \mu_1 + \eta_1 + 0.5\eta_0 = 0.5 + \mu_0 + \varepsilon_1 + \eta_1 + 0.5\eta_0$,
$y_2 = \mu_2 + \eta_2 + 0.5\eta_1 = (0.5 + \mu_1 + \varepsilon_2) + \eta_2 + 0.5\eta_1 = \mu_0 + 0.5 \cdot 2 + \varepsilon_2 + \varepsilon_1 + \eta_2 + 0.5\eta_1$,
$y_3 = \mu_3 + \eta_3 + 0.5\eta_2 = (0.5 + \mu_2 + \varepsilon_3) + \eta_3 + 0.5\eta_2 = \mu_0 + 0.5 \cdot 3 + \varepsilon_3 + \varepsilon_2 + \varepsilon_1 + \eta_3 + 0.5\eta_2$,
$y_4 = \mu_4 + \eta_4 + 0.5\eta_3 = (0.5 + \mu_3 + \varepsilon_4) + \eta_4 + 0.5\eta_3$
$= \mu_0 + 0.5 \cdot 4 + \varepsilon_4 + \varepsilon_3 + \varepsilon_2 + \varepsilon_1 + \eta_4 + 0.5\eta_3$,
…

$$y_t = \mu_0 + 0.5 \cdot t + \sum_{i=1}^{t} \varepsilon_i + \eta_t + 0.5\eta_{t-1}.$$

In period zero, the y_0 is: $y_0 = \mu_0 + \eta_0 + 0.5\eta_{-1}$. For given values of μ_0, η_0 and η_{-1}:

$$y_t = (y_0 - \eta_0 - 0.5\eta_{-1}) + 0.5 \cdot t + \sum_{i=1}^{t} \varepsilon_i + \eta_t + 0.5\eta_{t-1}.$$

Updating by s periods, the solution for y_{t+s} can be written as:

$$y_{t+s} = (y_t - \eta_t - 0.5\eta_{t-1}) + 0.5 \cdot s + \sum_{i=1}^{s} \varepsilon_{t+i} + \eta_{t+s} + 0.5\eta_{t+s-1}.$$

Taking the conditional expectation, the forecast function is:

$$\boxed{\begin{aligned} E_t y_{t+s} &= y_t - 0.5\eta_t - 0.5\eta_{t-1} + 0.5, \quad s=1 \\ E_t y_{t+s} &= y_t - \eta_t - 0.5\eta_{t-1} + 0.5 \cdot s, \quad s > 1 \end{aligned}}$$

Note that $\{y_t\}$ is a general trend plus irregular process. The change in the trend (i.e., $\Delta\mu_t$) consists of a stochastic component ε_t plus a deterministic component (i.e., change 0.5). The irregular component contains the stationary MA(1) process $\eta_t + 0.5\eta_{t-1}$. Other properties of this model can be seen in Section 8.

G. How can you make the models of Parts B and D stationary?

Answer: For Part B, it is possible to subtract y_{t-1} from each side of $y_t = 1.1\, y_{t-1} + \varepsilon_t$ to obtain: $\Delta y_t = 0.1 y_{t-1} + \varepsilon_t$. To examine the properties of the $\{\Delta y_t\}$ sequence, add and subtract $0.1 y_{t-2}$ to the right-hand side to obtain $\Delta y_t = 0.1\Delta y_{t-1} + 0.1 y_{t-2} + \varepsilon_t$. Now add and subtract $0.1 y_{t-3}$ to obtain: $\Delta y_t = 0.1\Delta y_{t-1} + 0.1\Delta y_{t-2} + 0.1 y_{t-3} + \varepsilon_t$. Continuing in this fashion yields $\Delta y_t = 0.1(\Delta y_{t-1} + \Delta y_{t-2} + \Delta y_{t-3} + \dots) + \varepsilon_t$. Obviously, the first differences are not stationary.

For Part D, write $y_t = y_{t-1} + t + \varepsilon_t$ as: $\Delta y_t = t + \varepsilon_t$. Now, simply detrend the Δy_t sequence to obtain a stationary sequence.

H. Does model E have an ARIMA(p,1,q) representation?

Answer: The model has an ARIMA(0,1,2) representation. Recall that: $y_t = \mu_t+\eta_t+0.5\eta_{t-1}$, where $\mu_t = \mu_{t-1}+\varepsilon_t$. Differencing once yields:
$\Delta y_t = \varepsilon_t + (1 - 0.5L - 0.5L^2)\eta_t$, where Δy_t is stationary.

The $\{\Delta y_t\}$ sequence is stationary as:

$E(\Delta y_t) = 0$, $\text{Var}(\Delta y_t) = \sigma^2 + 1.5\sigma_\eta^2$, and all autocorrelations are constant.

The first order autocorrelation of Δy_t is:

$$\rho_1 = \text{Cov}(\Delta y_t,\Delta y_{t-1}) / [\sqrt{\text{Var}(\Delta y_t)}\sqrt{\text{Var}(\Delta y_{t-1})}] = -0.25\sigma^2_\eta / (\sigma^2 + 1.5\sigma_\eta^2)$$

The second order autocorrelation of Δy_t is:

$$\rho_2 = \text{Cov}(\Delta y_t,\Delta y_{t-2}) /[\sqrt{\text{Var}(\Delta y_t)}\sqrt{\text{Var}(\Delta y_{t-2})}] = -0.5\sigma_\eta^2 / (\sigma^2+1.5\sigma_\eta^2)$$

The third order autocorrelation of Δy_t is:

$$\rho_3 = \text{Cov}(\Delta y_t,\Delta y_{t-3}) /[\sqrt{\text{Var}(\Delta y_t)}\sqrt{\text{Var}(\Delta y_{t-3})}] = -0 / (\sigma^2 + 1.5\sigma^2_\eta) = 0$$

and: $\rho_4 = \rho_5 = ... = 0$.

Thus, the autocorrelation function of Δy_t has the same characteristics as an MA(2) process, hence, y_t has an ARIMA(0,1,2) representation.

2. Given the initial condition y_0, find the general solution and the forecast function for the following variants of the trend plus irregular model:

A. $y_t = \mu_t+v_t$ where $\mu_t = \mu_{t-1}+\varepsilon_t$, $v_t = (1+\beta_1 L)\, \eta_t$ and $E\varepsilon_t\eta_t = 0$.

Answer: Iterate from period zero forward to obtain:

$$y_1 = \mu_1+v_1 = \mu_0+\varepsilon_1+\eta_1+\beta_1\eta_0,$$
$$y_2 = \mu_2+v_2 = \mu_1+\varepsilon_2+\eta_2+\beta_1\eta_1 = \mu_0+\varepsilon_2+\varepsilon_1+\eta_2+\beta_1\eta_1,$$
$$y_3 = \mu_3+v_3 = \mu_2+\varepsilon_3+\eta_3+\beta_1\eta_2 = \mu_0+\varepsilon_3+\varepsilon_2+\varepsilon_1+\eta_3+\beta_1\eta_2,$$

–

$$y_t = \mu_0+\sum_{i=1}^{t}\varepsilon_i+\eta_t+\beta_1\eta_{t-1}.$$

Since $y_0 = \mu_0+v_0 = \mu_0+\eta_0+\beta_1\eta_{-1}$, it follows that:

$$y_t = (y_0-v_0) + \sum_{i=1}^{t}\varepsilon_i+\eta_t+\beta_1\eta_{t-1}$$

Hence the general solution for y_t is:

$$\boxed{y_t = y_0 - \eta_0 - \beta_1\eta_{-1} + \sum_{i=1}^{t}\varepsilon_i + \eta_t + \beta_1\eta_{t-1}}$$

Next, update by s periods so that:

$$y_{t+s} = y_t - \eta_t - \beta_1\eta_{t-1} + \sum_{i=1}^{s}\varepsilon_{t+i} + \eta_{t+s} + \beta_1\eta_{t+s-1}$$

If we take the conditional expectation, the forecast function is:

$$\boxed{\begin{array}{l} E_t y_{t+s} = y_t - \eta_t - \beta_1\eta_{t-1} + \beta_1\eta_t \ \text{ for } s=1 \text{ and} \\ E_t y_{t+s} = y_t - \eta_t - \beta_1\eta_{t-1},\ s\geq 2. \end{array}}$$

B. $y_t = \mu_t + v_t$ where $\mu_t = \mu_{t-1} + \varepsilon_t$ and $v_t = (1+\beta_1 L)\eta_t$ and the correlation between ε_t and η_t equals unity.
Answer: The general solution and the forecast function are identical to those given in part A. The point is that the underlying model is not identified unless the correlation between the innovations is known.

C. Find the ARIMA representation of each model.
Answer: The first difference of the model from Part A is:

$$\Delta y_t = \varepsilon_t + [1+(\beta_1-1)\mathrm{L} - \beta_1 L^2]\eta_t.$$

Given that $E\varepsilon_t\eta_t = 0$, it follows that $E(\Delta y_t) = 0$ and:

$$\mathrm{Var}(\Delta y_t) = \sigma^2 + 2(\beta^2{}_1 - \beta_1 + 1)\sigma^2{}_\eta, \text{ where } \sigma^2 \equiv \mathrm{Var}(\varepsilon_t),\ \sigma^2{}_\eta \equiv \mathrm{Var}(\eta_t).$$

The first order autocorrelation of Δy_t is:

$$\begin{aligned} \rho_1 &= \mathrm{Cov}(\Delta y_t, \Delta y_{t-1}) \,/\, [\sqrt{\mathrm{Var}(\Delta y_t)} \cdot \sqrt{\mathrm{Var}(\Delta y_{t-1})}\ ] \\ &= -(1-\beta_1)^2\sigma^2{}_\eta \,/\, [\sigma^2 + 2(\beta^2{}_1 - \beta_1 + 1)\sigma^2{}_\eta] \end{aligned}$$

The second order autocorrelation of Δy_t is:

$$\rho_2 = -\beta_1\sigma^2{}_\eta \,/\, [\sigma^2 + 2(\beta^2{}_1 - \beta_1 + 1)\sigma^2{}_\eta]$$

and all higher autocorrelations are zero. Hence, the model acts as an ARIMA(0, 1, 2) process such that:

$$\boxed{\begin{array}{l} \rho_1 = -(1-\beta_1)^2\sigma^2{}_\eta \,/\, [\sigma^2 + 2(\beta^2{}_1 - \beta_1 + 1)\sigma^2{}_\eta] \\ \rho_2 = -\beta_1\sigma^2{}_\eta \,/\, [\sigma^2 + 2(\beta^2{}_1 - \beta_1 + 1)\sigma^2{}_\eta] \\ \text{and } \rho_i = 0 \text{ for } i > 2. \end{array}}$$

Difference the model of Part B to obtain:

$\Delta y_t = \varepsilon_t + [1+(\beta_1\text{-}1)L\text{-}\beta_1 L^2]\eta_t.$

Given that $E\varepsilon_t\eta_t = \sigma\sigma_\eta$ and $E(\Delta y_t) = 0$,

$\text{Var}(\Delta y_t) = \sigma^2+2(\beta^2{}_1\text{-}\beta_1+1)\sigma^2{}_\eta+2\sigma\sigma_\eta$, where $\sigma^2 \equiv \text{Var}(\varepsilon_t)$, $\sigma^2{}_\eta \equiv \text{Var}(\eta_t)$.
The first order autocorrelation of Δy_t is:

$$\rho_1 = [(\beta_1\text{-}1)\sigma\sigma_\eta\text{-}(1\text{-}\beta_1)^2\sigma^2{}_\eta] / [\sigma^2+2(\beta^2{}_1\text{-}\beta_1+1)\sigma^2{}_\eta+2\sigma\sigma_\eta]$$

The second order autocorrelation of Δy_t is:

$$\rho_2 = (\text{-}\beta_1\sigma\sigma_\eta\text{-}\beta_1\sigma^2{}_\eta) / [\sigma^2+2(\beta^2{}_1\text{-}\beta_1+1)\sigma^2{}_\eta+2\sigma\sigma_\eta]$$

and all higher autocorrelations are zero. Hence, the model acts as an ARIMA(0, 1, 2) process such that

$$\boxed{\begin{array}{l} \rho_1 = [(\beta_1\text{-}1)\sigma\sigma_\eta\text{-}(1\text{-}\beta_1)^2\sigma^2{}_\eta] / [\sigma^2+2(\beta^2{}_1\text{-}\beta_1+1)\sigma^2{}_\eta+2\sigma\sigma_\eta] \\ \rho_2 = (\text{-}\beta_1\sigma\sigma_\eta\text{-}\beta_1\sigma^2{}_\eta) / [\sigma^2+2(\beta^2{}_1\text{-}\beta_1+1)\sigma^2{}_\eta+2\sigma\sigma_\eta] \\ \text{and all } \rho_i = 0 \text{ for } i > 2. \end{array}}$$

3. The programming manual contains the program for RATS users. Even if you do not use RATS, the logic is virtually identical for any software package. Notice that John's results (see part c) are similar to those reported from a standard *t*-table. Additional replications would improve the accuracy of his results.

Sample Program for STATA Users

```
clear
set seed 12  /*this is to make the results reproducible*/

/*A:an experiment of tossing a coin and a tetrahedron*/
set obs 100
/*the two functions generate random integers in the interval [1,2] and [1,4] respectively. So their sum should be in [2,6]. The numbers are stored in the variable "num"*/
gen num = int(1+2*uniform()) + int(1+4*uniform())
table num

set obs 1000  /*do the experiment using 1000 replications*/
gen num1=int(1+2*uniform())+int(1+4*uniform())
table num1
```

Page 66: Models With Trend

```
/*B:replicate the Monte Carlo results for the t-dist. of an autoregressive coefficient*/
capture program drop mc1
program mc1,rclass
   version 8.0
   drop _all
   set obs 500  /*you may also change the sample size here, say 100*/
   gen a0 = 2
   gen epsilon = invnorm(uniform())  /*epsilon is a random variable distributed as N(0,1)*/
   gen y = a0 + epsilon  /*y is generated using the equation y = 2 + et*/
   gen time = _n
   tsset time
   reg y l.y  /*estimate the model under the alternative hypothesis*/
   return scalar tdist1 = (_b[l.y]/_se[l.y]) /*obtain the t-stat of a single estimation*/
end

/*run the program mc1 a thousand times to obtain the Monte Carlo critical values of the t-stat*/
simulate "mc1" tdist1 = r(tdist1), reps(1000)
summarize tdist1, detail
sort tdist1
display tdist1[50] "  " tdist1[950]  /*show the critical values at 10% significance level*/
display tdist1[25] "  " tdist1[975]  /*show the critical values at 5% significance level*/
display tdist1[5] "  " tdist1[995]  /*show the critical values at 1% significance level*/

/*C:perform another Monte Carlo experiment*/
capture program drop mc2
program mc2,rclass
    version 8.0
    drop _all
    set obs 100
    gen a0=2
    gen ep = invnorm(uniform())
    gen time = _n
    gen y = sum(a0) + sum(time) + sum(ep)  /*y is generated using the equation y(t)=a0+y(t-
1)+a2*t+et where a0=2 and a2=1. To see the results using different values of a0 and a2, change
'gen a0=2' above to another value of a0 and change this line to gen y = sum(a0) + sum(a2*time) +
sum(ep)*/
    tsset time
    reg d.y l.y time
    return scalar tdist2 = (_b[l.y]/_se[l.y])
end

simulate "mc2" tdist2 = r(tdist2), reps(2000)
summarize tdist2, detail
```

```
/*D:replicate Dickey-Fuller Tau-statistic(no constant or time trend)*/
capture program drop mc3
program mc3,rclass
    version 8.0
    drop _all
    set obs 100
    gen ep = invnorm(uniform())
    gen time = _n
    gen y = sum(ep) /*y is generated using the equation y(t)=y(t-1)+ep*/
    tsset time
    reg d.y l.y,noconstant
    return scalar tdist3 = (_b[l.y]/_se[l.y])
end

simulate "mc3" tdist3 = r(tdist3), reps(2000)
summarize tdist3, detail
sort tdist3
display  tdist3[20] " "  tdist3[50] " "  tdist3[100] " "  tdist3[200]
```

4. The columns in the file labeled REAL.XLS contain the logarithm of the real exchange rates for Canada, Japan, Germany, and the U.K. The following program can be used to answer Question 4.

Sample Program for RATS Users

```
calendar 1973 2 12              ;*  Set the calendar for monthly data beginning
allocate 1989:12                ;*    with 1973:2
open data a:\real.xls           ;*  It is assumed that the data is on drive a:\
data(format=xls,org=obs) /
table                           ;*  Create the sample statistics of each series

;*  The following steps can be repeated for each real exchange rate series.
dif rcan / drcan          ;*  Create the first-difference of the Canadian rate.

* Now find the ACF of rcan and drcan
cor rcan
cor drcan
* Notice the slow decay of the autocorrelations of rcan. This can be indicative of a unit root.

* Perform the Dickey-Fuller tests
lin drcan ; # constant rcan{1}
lin drcan ;                      ;*  Now use 12 lags
# constant rcan{1} drcan{1 to 12}
```
* Since neither have a trend, use the τ_μ statistic

```
set trend = t                                          ;* create a time trend
lin drcan ; # constant rcan{1} drcan{1 to 12}  ;*  Now use 12 lags plus the time trend
```

* Use the τ_τ statistic to test for a unit root

Sample Program for STATA Users

```
clear
cd "x:\New_data" /*change this line to wherever the dataset is saved*/
insheet using real.txt
gen obs = _n
tsset obs /*declare this is a time series dataset*/

/*A:find the ACF and PACF*/
corrgram rcan,lag(12) /*ACF and PACF of the level of the real exchange rate*/
corrgram rger,lag(12)
corrgram rjap,lag(12)
corrgram ruk,lag(12)

/*ACF and PACF of the first difference of the real exchange rate*/
corrgram d.rcan, lag(12)
corrgram d.rger,lag(12)
corrgram d.rjap,lag(12)
corrgram d.ruk,lag(12)

/*ACF and PACF of the detrended real exchange rate*/
reg rcan obs   /*first regress the real exchange rate on trend*/
/*obtain the residual series, which is the detrended real exchange rate*/
predict residrcan,residuals
corrgram residrcan, lag(12)

reg rger obs
predict residrger,residuals
corrgram residrger, lag(12)

reg rjap obs
predict residrjap,residuals
corrgram residrjap, lag(12)

reg ruk obs
predict residruk,residuals
corrgram residruk, lag(12)

/*C:DF and ADF tests of unit root*/
dfuller rcan                  /*Dicky-Fuller test for RCAN (No lags)*/
```

```
dfuller rcan, lags(12)        /*Augmented Dicky-Fuller test for RCAN (12 lags)*/
dfuller rcan, lags(12) trend  /*Augmented Dicky-Fuller test for RCAN (Trend + 12
lags)*/

dfuller rger
dfuller rger, lags(12)
dfuller rger, lags(12) trend

dfuller rjap
dfuller rjap, lags(12)
dfuller rjap, lags(12) trend

dfuller ruk
dfuller ruk, lags(12)
dfuller ruk, lags(12) trend
```

5. The seasonal unit root procedures can be downloaded from the Estima website www.estima.com. HEGYQNEW.SRC performs the Hylleberg, Engle, Granger, and Yoo (HEGY) unit root test on quarterly data. MHEGY.SRC performs the test using monthly data.

```
cal 1960 1 4                 ;* Read in the data set. The format is the same as that for Chapter 2
all 2002:1
open data a:\quarterly.xls
data(format=xls,org=obs)

log m1nsa / lm1                          ;* Call lm1 the log of the M1NSA series.

source(noecho) c:\winrats\myhegy.src     ;* Compile the procedure
* Now implement the procedure. The maximum number of lags in 8. The lag length is chosen by
* removing lags that are not significant that the 5% level.
@hegyqnew(signif=0.05,criterion=signlag,nlags=8) lm1

* Alternatively, you can make your own transformations as follows:
dif lm1 / dlm1                           ;* Difference lm1
dif(sdiffs=1) lm1 / d4lm1                ;* Create the seasonal difference
set trend = t                            ;* Create a time trend
set y1 = lm1+lm1{1}+lm1{2}+lm1{3}        ; Create y1t
set y2 = lm1-lm1{1}+lm1{2}-lm1{3}        ;* Create y2t
set y3 = lm1 - lm1{2}                    ;* Create y3t

* Now regress the seasonal difference on a constant, trend, y1t-1, y2t-1, y3t-1, y3t-2, and various lags.
* You can experiment with different lag lengths and by including the deterministic seasonal
* dummy variables.
```

Page 70: Models With Trend

```
lin d4lm1
# constant trend y1{1} y2{1} y3{1} y3{2} d4lm1{1 to 2}
```

Program for STATA Users

```
clear
cd "x:\New_data" /*change this line to wherever the dataset is saved*/
insheet using quarterly.txt
gen time = q(1960q1)+_n-1 /*the dataset starts from 1960Q1*/
tsset time,format(%tq) /*declare this is a quarterly timeseries dataset*/

gen yt = ln(m1nsa)
/*form the four variables for quarterly data*/
gen y1t = l.yt + l2.yt + l3.yt + l4.yt
gen y2t = l.yt - l2.yt + l3.yt - l4.yt
gen y3t = l.yt - l3.yt
gen y4t = l2.yt - l4.yt
/*estimate the regression(Intercept + Seasonal dummies + time + no lags)*/
reg s4.yt time y1t y2t y3t y4t
test y3t y4t  /*determine whether there is a seasonal unit root in M1*/

/*use the package "hegy4" to perform the test for seasonal unit root in a quarterly timeseries.
First search "seasonal unit root" in the range "all",then click to install the module "HEGY4"*/
hegy4 yt, lags(1/9) det(strend) /*estimate the regression(Intercept + Seasonal dummies + time
+ 9 lags)*/
```

6. The second column in the file BREAK.XLS contains the simulated data used in Section 8.

RATS PROGRAM FOR PARTS A and B

The data set contains 100 simulated observations with a break occurring at t = 51. To replicate the results in section 8, perform the following:

```
all 100                          ;* These three lines read in the data set
open data a:\break.xls
data(format=xls,org=obs)

set trend = t                    ;* Creates a time trend

* The graph of the series shown in Figure 4.10 was created using
graph 1 ; # y1

* The ACF is obtained using
cor(method=yule) y1
```

```
dif y1 / dy1                    ;* Difference y1
* The next three lines perform the Dickey-Fuller tests reported on the top of page 207.
linreg dy1 ; # y1{1}
linreg dy1 ; # constant y1{1}
linreg dy1 ; # constant trend y1{1}
* Next, create a dummy variable with a break occurring in period 51
set dummy = %if(t.ge.51,1,0)
```

* Regress y1 on a constant, a trend and the dummy variable. Save the residuals as *resids*

```
lin y1 / resids ; # constant trend dummy
```

* Regress *resids* on its own first lag

```
linreg resids ; # resids{1}
```

Continuation for parts C through G

```
graph 1 ; # y2           ;* Create the graph of y2

cor y2                   ;* Creates the ACF of y2. Note that the ACF shows little decay
dif y2 / dy2             ;* Difference y2

* Now perform the Dickey-Fuller test. Diagnostic testing does not reveal the need for
* any lagged changes. Nevertheless, the constant and trend do not appear to be significant.
lin dy2 ; # constant trend y2{1}
```

* Now estimate $y2_t = a_0 + a_2t + \mu_2D_L$ and save the residuals as *resids*

```
lin y2 / resids; # constant trend dummy
```

* Regress *resids* on its first lag

```
lin resids ; # resids{1}

* The coefficient on the lagged residual is 0.965. The series seems to contain a unit root.
```

* Instead create a pulse dummy called *d*2; *d*2 = 1 only in period 51

```
set d2 = %if(t.eq.51,1,0)
```

* Regress y*2* on a constant, a trend and *d*2 and save the residuals as *resids*

```
lin y2 / resids; # constant trend d2
```

* Regress *resids* on its first lag

```
lin resids ; # resids{1}

* The coefficient on the lagged residual is 0.982. Hence, there is a unit root.
```

Sample Program for Stata Users

```
clear
cd "x:\New_data"
insheet using quarterly.txt
gen obs = _n
tsset obs

/*A:replicate the result of section 9*/
gen lgdp=ln(gdp)
/*estimate the logarithmic first difference of real gdp as an AR(2) process*/
reg d.lgdp l.(d.lgdp) l2.(d.lgdp)

dfuller lgdp,trend lags(2) regress  /*replicate the result of (4.45)*/
dfuller lgdp,lags(2) regress        /*replicate the result of (4.46)*/

/*B:perform the Perron test to determine if real gdp is trend stationary with a break occuring
in mid-1973*/
gen x=0
replace x=1 if _n>=55  /*generate a dummy variable x which takes the value 1 after
1973Q3*/
reg lgdp x obs
predict resid,residuals
dfuller resid,noconstant
reg resid l.resid, noconstant
predict resid1,residuals
corrgram resid1,lag(12)
```

7. The file QUARTERLY.XLS contains the real GDP data that was used to estimate (4.45).

A. The following program will replicate the results in Section 9.

* The data set runs from 1960Q1 through 2002Q1. Read in the data with

```
cal 1960 1 4
all 2002:1
open data a:\quarterly.xls
data(format=xls,org=obs)

set time = t                  ; * Create a trend
set ly = log(gdp)             ; * Create ly as the log of GDP
dif ly / dly                  ; * Create the first difference of ly

* You can estimate (4.42) using
lin dly; # constant time ly{1} dly{1 to 2}
```

```
exc ; # time ly{1}                        ;* Test for the drift using the φ3 statistic
exc ; # constant time ly{1}               ;* Test for the constant and trend using φ2

* Estimate the model without the trend
lin dly; # constant ly{1} dly{1 to 2}
exc ; # constant ly{1}                    ;* Test for the drift using the φ1 statistic
```

B. Create the level and pulse dummies beginning with the fourth quarter of 1973

```
set level = %if(t.ge.1973:4,1,0)
set pulse
```

C. * RATS contains the procedures BNDECOMP.SRC and HPFILTER.SRC that allow you to perform the Beveridge-Nelson decomnposition and the HP filter. If you do not have the procedures, you can download them from the Estima website www.estima.com. You can perform the two decompositions using

```
source(noecho) c:\winrats\bndecomp.src
@bndecomp(log) gdp perm  1960:1 1970:1 2002:1 2 0
source(noecho) c:\winrats\hpfilter.src
@hpfilter gdp / hp_rgdp
```

Sample Program for STATA Users

```
clear
cd "x:\New_data"
insheet using quarterly.txt
gen obs = _n
tsset obs

/*A:replicate the result of section 9*/
gen lgdp=ln(gdp)
/*estimate the logarithmic first difference of real gdp as an AR(2) process*/
reg d.lgdp l.(d.lgdp) l2.(d.lgdp)

dfuller lgdp,trend lags(2) regress  /*replicate the result of (4.45)*/
dfuller lgdp,lags(2) regress        /*replicate the result of (4.46)*/

/*B:perform the Perron test to determine if real gdp is trend stationary with a break occuring
in mid-1973*/
gen x=0
replace x=1 if _n>=55  /*generate a dummy variable x which takes the value 1 after 1973Q3*/
reg lgdp x obs
predict resid,residuals
```

```
dfuller resid,noconstant
reg resid l.resid, noconstant
predict resid1,residuals
corrgram resid1,lag(12)
```

8. The file PANEL.XLS contains the real exchange rate series used to perform the panel unit root tests reported in Section 12.

   A. Replicate the results of Section 12.

SAMPLE RATS PROGRAM

```
cal 1980 1 4                ;* The first four lines read in the data set. The quarterly observations
all 2002:2                  ;* begin in 1980Q1 and end in 2002Q2.
open data a:\panel.xls
data(format=xls,org=obs)

* To select lag lengths, you can begin with 8 lags and pare down the model by dropping a lag if
* it is not significant at the 5% level. An automated way to do this is to use a do loop. Consider

log australia / lx ; dif lx / dlx     ;* Form the logarithmic change in the real Australian rate
do i = 8,1,-1                         ;* Use lag 8 through 1 in steps of -1
  * Now, perform a augmented Dickey-Fuller test  using lags 1 to i
  lin(noprint) dlx ; # dlx{1 to i} lx{1} constant

  dis i %tstats(i) %tstats(i+1)     ; Display the t-stat for lag i and for Dickey-Fuller test
end do i

*Now repeat for Canada
log canada / lx ; dif lx / dlx
do i = 8,1,-1
     lin(noprint) dlx ; # dlx{1 to i} lx{1} constant
     dis i %tstats(i) %beta(i+1) %tstats(i+1)
end do i

* Repeat for each country for find the lag lengths reported in Table 4.8.
* To find the common time effect, form the average value of the real rate

set aver = (australia+canada+france+germany+japan+netherlands+UK+US)/8

* Subtract the time average from each rate. Since we are using logs, we can use the following
* to create log of the modified Australian real rate
set lx = log(australia/aver); dif lx / dlx

* Repeat for each nation and perform the Dickey-Fuller test for each. For the Australian case
do i = 8,1,-1
```

```
        lin(noprint) dlx ; # dlx{1 to i} lx{1} constant
        dis i %tstats(i) %beta(i+1) %tstats(i+1)
    end do i
```

B. If Germany and Canada are excluded from the sample, the evidence in favor of PPP is much stronger. However, the selection of the panel should be made on theoretical grounds.

C. As in a standard unit root test, the power of a panel unit root test is decreased in the presence of a time trend that is not actually in the data generating process.

Q9. The Programming Manual contains the complete answer Questions 9 for RATS users.

Sample Program for STATA Users

```
clear
set seed 12

/*A:an experiment of tossing two tetrahedra*/
set obs 1000  /*there are 1000 trials of the experiment*/
gen num=int(1+4*uniform())+int(1+4*uniform())
table num

/*B:show the downward bias for the first-order AR coefficient in sample of 100 and 50
observations*/
/*this is to replicate the results with a sample size of 100*/
capture program drop bias
program define bias,rclass
   version 8.0
   drop _all
   args alpha1
   set obs 100 /*to see the results of a smaller sample size, just change the obs to 50*/
   gen time = _n
   tsset time
   gen y=0
   gen ep=invnorm(uniform())
   forvalues i=1/100 {
          replace y=`alpha1'*y[`i']+ep[`i'+1] if (_n>`i')
   }
   reg y l.y
   return scalar discrep = `alpha1' - _b[l.y]
end

simulate "bias 0.2" discrep = r(discrep),reps(1000)
summarize discrep
simulate "bias 0.5" discrep = r(discrep),reps(1000)
```

```
summarize discrep
simulate "bias 0.9" discrep = r(discrep),reps(1000)
summarize discrep
simulate "bias 0.99" discrep = r(discrep),reps(1000)
summarize discrep
simulate "bias 0" discrep = r(discrep),reps(1000)
summarize discrep

/*C:Dickey-Fuller critical values(constant + time trend)*/
capture program drop dfcv
program dfcv,rclass
    version 8.0
    drop _all
    set obs 100
    gen ep = invnorm(uniform())
    gen y = sum(ep)
    gen time = _n
    tsset time
    reg d.y l.y time /*to obtain the DF critical values for a model with constant but no time
trend, just drop the time variable here*/
    return scalar contime = (_b[l.y]/_se[l.y])
end

simulate "dfcv" contime = r(contime), reps(10000)
sort contime
/*show the critical values of the Tau(tau) statistic at 1%,2.5%,5%,10% significance levels*/
display contime[100] "  " contime[250] "  " contime[500] "  " contime[1000]

/*D:power of the DF test(Constant but No Time Trend)*/
capture program drop power
program define power,rclass
   version 8.0
   args alpha1
   drop _all
   set obs 100 /*to see the results of a smaller sample size, just change the obs to 50*/
   gen time = _n
   tsset time
   gen y=0
   gen ep=invnorm(uniform())
   scalar hits10=0
   scalar hits5=0
   scalar hits1=0
   local i=1
   while `i'<=100 {
```

```
            replace y = `alpha1'*y[`i'] + sqrt(1-(`alpha1')^2)*ep[`i'+1] if (_n>`i')
            local i=`i'+1
            }
   reg d.y l.y  /*to obtain the power of the DF stat(Constant+Time Trend), just include
"time" variable in this regression*/
         if abs( _b[l.y]/_se[l.y])>2.58 {
            return scalar hits10 = hits10+1
           }
         if abs( _b[l.y]/_se[l.y])>2.89 {
            return scalar hits5 = hits5+1
           }
         if abs( _b[l.y]/_se[l.y])>3.51 {
            return scalar hits1 = hits1+1
           }

end

simulate "power 0.8" hits10 = r(hits10) hits5 = r(hits5) hits1 = r(hits1),reps(1000)
summarize hits10
display " when alpha1=0.8,the power of the test(10% significance level) = "=r(N)
summarize hits5
display " when alpha1=0.8,the power of the test(5% significance level)= "=r(N)
summarize hits1
display " when alpha1=0.8,the power of the test(1% significance level)= "=r(N)

simulate "power 0.9" hits10 = r(hits10) hits5 = r(hits5) hits1 = r(hits1),reps(1000)
summarize hits10
display " when alpha1=0.9,the power of the test(10% significance level) = "=r(N)
summarize hits5
display " when alpha1=0.9,the power of the test(5% significance level)= "=r(N)
summarize hits1
display " when alpha1=0.9,the power of the test(1% significance level)= "=r(N)

simulate "power 0.95" hits10 = r(hits10) hits5 = r(hits5) hits1 = r(hits1),reps(1000)
summarize hits10
display " when alpha1=0.95,the power of the test(10% significance level) = "=r(N)
summarize hits5
display " when alpha1=0.95,the power of the test(5% significance level)= "=r(N)
summarize hits1
display " when alpha1=0.95,the power of the test(1% significance level)= "=r(N)

simulate "power 0.99" hits10 = r(hits10) hits5 = r(hits5) hits1 = r(hits1),reps(1000)
summarize hits10
```

```
display " when alpha1=0.99,the power of the test(10% significance level) = "=r(N)
summarize hits5
display " when alpha1=0.99,the power of the test(5% significance level)= "=r(N)
summarize hits1
display " when alpha1=0.99,the power of the test(1% significance level)= "=r(N)
```

10. Section 7 of Chapter 4 of the Programming Manual contains a discussion of bootstrapping. In particular, Program 4.14 provides an example of bootstrapping regression coefficients.

A. Use the data in MONEYDEM.XLS to estimate the logarithmic change in real GDP ($dlrgdp_t$) as an AR(2) process. You should obtain

$$dlrgdp_t = 0.005 + 0.0251 dlrgdp_{t-1} + 0.136 dlrgdp_{t-2} + \varepsilon_t$$

As shown in the program, construct the 95% confidence interval for the AR(2) coefficient 0.0136.

B. Follow the program to obtain the bootstrap 95% confidence interval for the coefficient.

C. How would you modify the program so as to obtain the bootstrap 95% confidence interval for the AR(1) coefficient?

Sample Program for STATA Users

```
/*obtain the bootstrap 95% confidence interval for the AR(2) and AR(1) coefficients.*/
clear
cd "X:\New_data" /*Please change the directory to wherever the dataset is saved*/
insheet using money_dem.txt
gen obs = _n
tsset obs
gen lrgdp = ln(rgdp)
gen dlrgdp = d.lrgdp
reg dlrgdp l.dlrgdp l2.dlrgdp  /*estimate dlrgdp as an AR(2) process*/
capture save moneydem,replace
predict resids, residuals
keep resids
/*"resids" is the series from which the bootstrapped residuals are drawn */
capture save resids,replace

scalar beta1hat = _b[L.dlrgdp]
scalar beta2hat = _b[L2.dlrgdp]
scalar beta0hat = _b[_cons]

capture program drop bsbeta
program define bsbeta,rclass
   version 8.0
   use resids,clear
```

```
    gen rn = 4 + int((_N-4+1)*uniform()) /*the first three observations of dlrgdp are missing
due to the first difference and the two lags of rgdp. Hence the sample is drawn in [4,_N] or
[4,169]*/
    gen estar = resids[rn]  /*obtain the bootstrapped residuals*/
    merge using moneydem
    tsset obs
    scalar ii = 2 + int((_N-2)*uniform()) /*the first observation of dlrgdp is missing due to the
first difference of rgdp. Moreover, the upper bound should be _N-1 to ensure the second
initial value is not missing. Hence "ii" is drawn in [2,_N) or [2,169)*/
    gen ystar = dlrgdp[ii]
    replace ystar = dlrgdp[ii+1] if _n > 1
    forvalues i = 1/167 {
     replace ystar = beta1hat*ystar[`i'+1] + beta2hat*ystar[`i'] + beta0hat + estar[`i'] if _n>`i'+1
          }
    reg ystar l.ystar l2.ystar  /*estimate the model using the bootstrapped sample*/
    return scalar beta2star = _b[L2.ystar]
    return scalar beta1star = _b[L.ystar]
end

set seed 2001
simulate "bsbeta" beta2star = r(beta2star) beta1star = r(beta1star),reps(1000)
summarize beta2star,detail
sort beta2star
display "90% confidence interval for beta2:"  beta2star[50] "  " beta2star[950]
display "95% confidence interval for beta2:"  beta2star[25] "  " beta2star[975]
display "95% confidence interval for beta2:"  beta2star[5] "  " beta2star[995]
/*obtain the bootstrap 95% confidence interval for the AR(1) coefficient*/
quietly summarize beta1star,detail
sort beta1star
display "95% confidence interval for beta2:"  beta1star[25] "  " beta1star[975]
```

CHAPTER 5

MULTIEQUATION TIME-SERIES MODELS

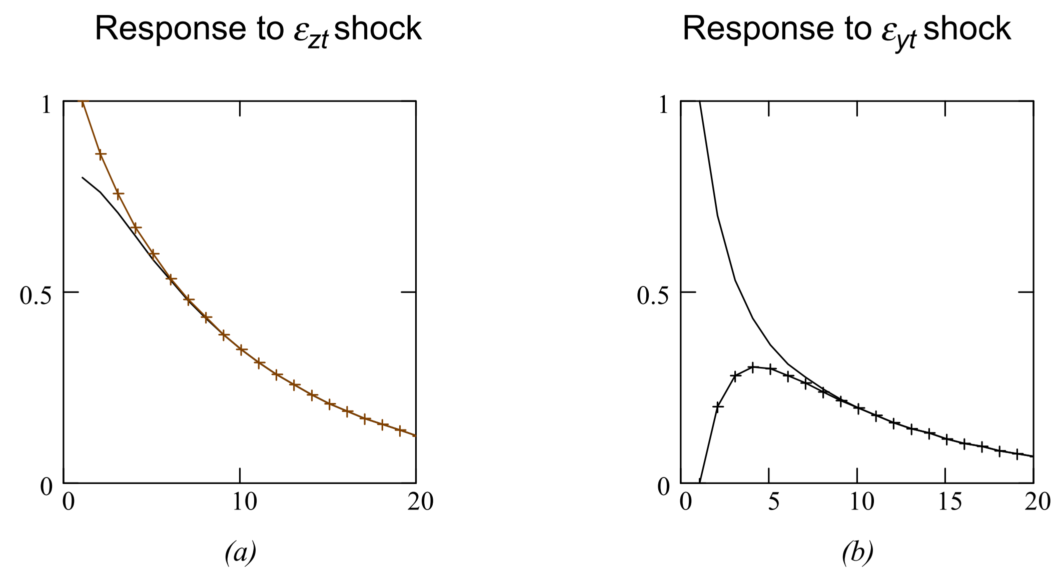
Lecture Suggestions

1. Although many economists are skeptical of parsimonious intervention and transfer function models, it is wise not to skip Sections 1 through 3. These sections act as an introduction to VAR analysis. In a sense, VAR analysis can be viewed as a progression. Intervention analysis treats $\{y_t\}$ as stochastic and $\{z_t\}$ as a deterministic process. Transfer function allows $\{z_t\}$ to be stochastic, but assumes that there is no feedback from the $\{y_t\}$ sequence to the $\{z_t\}$ sequence. The notion of an autoregressive distributed lag (ADL) is also introduced here. Finally, VAR analysis treats all variables symmetrically. Use Section 4 to explain the limitations of intervention and transfer function analysis and to justify Sims' methodology.

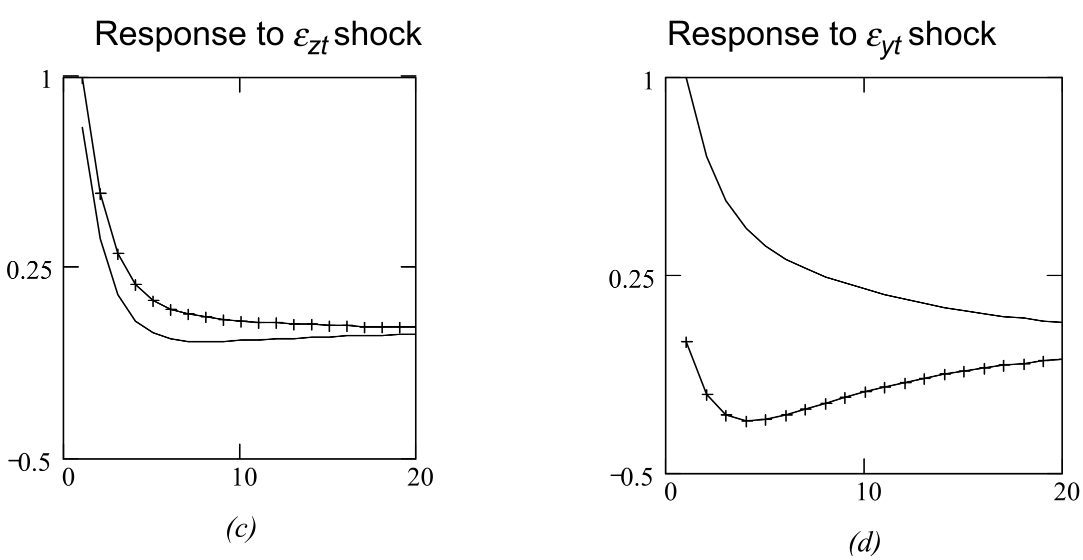
2. You should emphasize the distinction between the VAR residuals and the structural innovations. Questions 4, 5, and 6 at the end of the chapter are especially important. I work through one of these questions in the classroom and assign the other two for homework. You might also make an overhead of Figure 5.7 in order to illustrate the effects of alternative orderings in a Choleski decomposition. A large-sized version of the figure is included here for your convenience.

Figure 5.7: Two Impulse Response Functions

Model 1: $$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} 0.7 & 0.2 \\ 0.2 & 0.7 \end{pmatrix} \cdot \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} u_t \\ v_t \end{pmatrix}$$



Model 2: $$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} 0.7 & -0.2 \\ -0.2 & 0.7 \end{pmatrix} \cdot \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} u_t \\ v_t \end{pmatrix}$$



Legend: Solid line = $\{y_t\}$ sequence Cross-hatch = $\{z_t\}$ sequence

Note: In all cases $u_t = 0.8v_t + \varepsilon_{yt}$ and $v_t = \varepsilon_{zt}$

Answers to Questions

1. Consider three forms of the intervention variable:

    pulse: $z_1 = 1$ and all other $z_i = 0$

    pure jump: $z_1 = z_2 = ... = 1$ and all other $z_i = 0$

    prolonged impulse: $z_1 = 1$; $z_2 = 0.75$; $z_3 = 0.5$; $z_4 = 0.25$; and all other values of $z_i = 0$

A. Show how each of the following $\{y_t\}$ sequences responds to the three types of interventions:

i) $y_t = 0.5y_{t-1} + z_t + \varepsilon_t$

Answer: Using lag operators to solve for y_t:

$$y_t = (z_t + \varepsilon_t)/(1 - 0.5L)$$
$$= (1 + 0.5L + 0.25L^2 + 0.125L^3 + ...)z_t + \varepsilon_t/(1 - 0.5L)$$

Note that $dy_{t+i}/dz_t = dy_t/dz_{t-i}$ and that for a pure pulse, only $z_1 = 1$. Hence:

$dy_t/dz_t = 1$; $dy_{t+1}/dz_t = dy_t/dz_{t-1} = 0.5$; $dy_{t+2}/dz_t = dy_t/dz_{t-2} = 0.25$; ... . Thus, for $t \geq 1$:

$$\boxed{dy_t/dz_1 = (0.5)^{t-1}}$$

For the pure jump, $z_1 = z_2 = ... = 1$ and all other $z_i = 0$ so that:

$dy_1/dz_1 = 1$; $dy_2/dz_1 = 1 + 0.5$; $dy_3/dz_1 = 1 + 0.5 + 0.25$; ... . Thus, for the pure jump:

$$\boxed{dy_t/dz_1 = 1 + 0.5 + 0.5^2 + 0.5^3 + ... + 0.5^{t-1}}$$

Notice that as $t \to\infty$, the limiting value of the pure jump is: $dy_t/dz_1 = 1/(1\text{-}0.5) = 2.0$. For the prolonged impulse, use the solution for y_t and treat: $dz_1 = 1$; $dz_2/dz_1 = 0.75$; $dz_3/dz_1 = 0.5$; $dz_4/dz_1 = 0.25$ and all other values of $dz_i = 0$. Thus, the impulse responses are: obtained as:

$dy_1/dz_1 = 1$; $dy_2/dz_1 = 1(0.75) + 0.5(1)$; $dy_3/dz_1 = 1(0.5)+0.5(0.75) + 0.25(1)$; and for $t \geq 4$:

$$\boxed{dy_t/dz_1 = 0.5^{t-1} + 0.75(0.5)^{t-2} + 0.5(0.5)^{t-3} + 0.25(0.5)^{t-4}}$$

ii) $y_t = -0.5y_{t-1} + z_t + \varepsilon_t$

Answer: Use lag operators to write:

$(1 + 0.5L)\, y_t = z_t + \varepsilon_t$.

Explicitly solving for y_t:

$$y_t = (z_t + \varepsilon_t)/(1+0.5L)$$
$$= z_t + (-0.5)z_{t-1} + (-0.5)^2 z_{t-2} + (-0.5)^3 z_{t-3} + ... + (-0.5)^i z_{t-i} + ... + \varepsilon_t/(1+0.5L)$$

For the pulse intervention: $z_1 = 1$ and all other $z_i = 0$. Hence:

$dy_1/dz_1 = 1$, $dy_2/dz_1 = -0.5$, $dy_3/dz_1 = (-0.5)^2$, ... Thus, for $t \geq 1$:

$$\boxed{dy_t/dz_1 = (-0.5)^{t-1}}$$

For the pure jump: $z_1 = z_2 = ... = 1$. Hence:

$dy_1/dz_1 = 1$, $dy_2/dz_1 = 1 + (-0.5)$, $dy_3/dz_1 = 1 + (-0.5) + (-0.5)^2$, ... Thus, for $t \geq 1$:

$$\boxed{dy_t/dz_1 = 1 + (-0.5) + (-0.5)^2 + ... + (-0.5)^{t-1}}$$

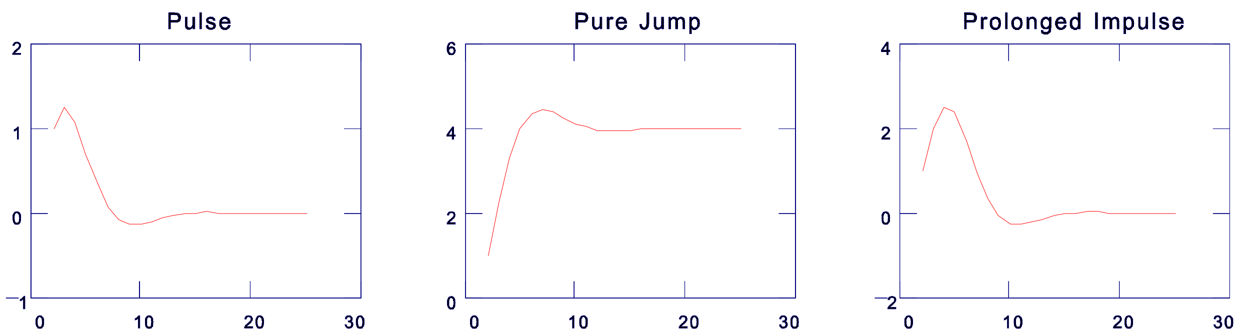
For the prolonged impulse: $z_1 = 1$; $z_2 = 0.75$; $z_3 = 0.5$; $z_4 = 0.25$; and all other values of $z_i = 0$. Hence:

$dy_1/dz_1 = 1$, $dy_2/dz_1 = 0.75 + (-0.5) \cdot 1$, $dy_3/dz_1 = 0.5 + 0.75 \cdot (-0.5) + 1 \cdot (-0.5)^2$,
$dy_4/dz_1 = 0.25 + 0.5 \cdot (-0.5) + 0.75\ (-0.5)^2 + 1\ (-0.5)^3$,
$dy_5/dz_1 = 0.25 \cdot (-0.5) + 0.5 \cdot (-0.5)^2 + 0.75 \cdot (-0.5)^3 + (-0.5)^4$, ... So that, for all $t \geq 4$:

$$\boxed{dy_t/dz_1 = 0.25\ (-0.5)^{t-4} + 0.5 \cdot (-0.5)^{t-3} + 0.75 \cdot (-0.5)^{t-2} + 1 \cdot (-0.5)^{t-1}}$$

iii) $y_t = 1.25y_{t-1} - 0.5y_{t-2} + z_t + \varepsilon_t$

Answer:



iv) $y_t = y_{t-1} + z_t + \varepsilon_t$

Answer: Write the solution for y_t using $(1-L)y_t = z_t + \varepsilon_t$, so that given y_0:

$$y_t = z_t + z_{t-1} + z_{t-2} + z_{t-3} + ... + z_1 + \varepsilon_t + ... + \varepsilon_1 + y_0$$

For the pulse, set $z_1 = 1$ and all other $z_i = 0$ so that:

$dy_1/dz_1 = 1$, $dy_2/dz_1 = 1$, $dy_3/dz_1 = 1$, ... so that for all $t \geq 1$:

$$\boxed{dy_t/dz_1 = 1}$$

For the pure jump, set $z_1 = z_2 = ... = 1$ and all other $z_i = 0$, so that:

$dy_1/dz_1 = 1$, $dy_2/dz_1 = 2$, $dy_3/dz_1 = 3$, ... Hence, for all $t \geq 1$:

$$\boxed{dy_t/dz_1 = t}$$

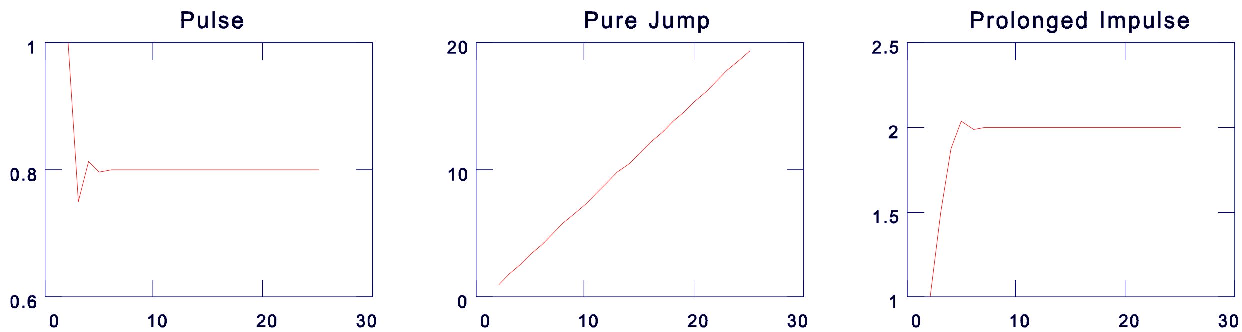
For the prolonged impulse set $z_1 = 1$; $z_2 = 0.75$; $z_3 = 0.5$; $z_4 = 0.25$; and all other values of $z_i = 0$ so that:

$dy_1/dz_1 = 1$, $dy_2/dz_1 = 0.75{+}1$, $dy_3/dz_1 = 0.5{+}0.75{+}1$, $dy_4/dz_1 = .25{+}0.5{+}0.75{+}1$, ... and for all $t \geq 4$:

$$\boxed{dy_t/dz_1 = 0.25+0.5+0.75+1}$$

v) $y_t = 0.75y_{t-1} + 0.25y_{t-2} + z_t + \varepsilon_t$

Answer: The time paths of the responses are given in the three figures below. The actual coefficients are derived in Part B.



B. Notice that the intervention models in iv) and v) have unit roots. Show that the intervention variable $z_1 = 1$; $z_2 = -1$; and all other values of $z_i = 0$, has only a temporary effect on these two sequences.

Answer: From the solution for y_t in part iv), $y_t = z_t + z_{t-1} + z_{t-2} + z_{t-3} + ... + z_1 + \varepsilon_t + ... + \varepsilon_1 + y_0$. Set $z_1 = 1$ and $z_2 = -1$ so that:

$dy_1/dz_1 = 1$, $dy_2/dz_1 = 1 - 1 = 0$, so that for all $t \geq 2$:

$$\boxed{dy_t/dz_1 = 0}$$

For the model in part v), $y_t = (z_t + \varepsilon_t)/[(1-L)(1+0.25L)]$. Abstracting from terms containing elements of the $\{\varepsilon_t\}$ sequence, the *improper* form of the particular solution is:

$$\begin{aligned} y_t &= (z_t + z_{t-1} + z_{t-2} + ...)/(1+0.25L) \\ &= (1 - 0.25L + 0.25^2L^2 - 0.25^3L^3 + ...)z_t + (1 - 0.25L + 0.25^2L^2 - 0.25^3L^3 + ...)z_{t-1} \\ &\quad + (1 - 0.25L + 0.25^2L^2 - 0.25^3L^3 + ...)z_{t-2} + ... \\ &= z_t + (1 - 0.25)z_{t-1} + (1 - 0.25 + 0.25^2)z_{t-2} + (1 - 0.25 + 0.25^2 - 0.25^3)z_{t-3} + ... \end{aligned}$$

To find the proper form of the solution, it is necessary to impose two initial conditions. However, from the improper form of the solution, it is clear that for $z_1 = 1$ and $z_2 = -1$:

$dy_1/dz_1 = 1$, $dy_2/dz_1 = -1 + 0.75 = -0.25$, $dy_3/dz_1 = 0 - 0.75 + (1 - 0.25 + 0.25^2)$, ... and:

$$\boxed{dy_t/dz_1 = (-0.25)^{t-1} \text{ for } t \geq 1}.$$

Hence, the long-run effect converges to zero.

C. Show that an intervention variable will not have a permanent effect on a unit root process if all values of z_i sum to zero.

Answer: If y_t has a single unit root, the model can be written in the form $(1-L)A(L)y_t = z_t + \varepsilon_t$.

To sketch the proof, abstract from the $\{\varepsilon_t\}$ sequence and consider the particular solution for y_t. Since there is a single unit root, all characteristic roots of $A(L)$ lie outside the unit circle and the expression $A(L)y_t$ is stationary. Now consider $z_t/(1\text{-}L) = z_t + z_{t-1} + z_{t-2} + \ldots$; by assumption this sum is zero. Beginning from a given value of y_0, the long-run effects of $\{z_t\}$ on $A(L)y_t$ (and on y_t itself) must converge toward zero.

D. Discuss the plausible models you might choose if the $\{y_t\}$ sequence is:

i) stationary and you suspect that the intervention has a permanent effect on Ey_t.
Answer: The intervention variable $\{z_t\}$ might be a pure jump or gradually increasing (decreasing) process.

ii) stationary and you suspect that the intervention has a growing and then a diminishing effect.
Answer: The intervention variable $\{z_t\}$ might be a prolonged impulse process.

iii) non-stationary and you suspect that the intervention has a permanent effect on the level of $\{y_t\}$.
Answer: Model $\{y_t\}$ as a unit root process and model the intervention variable $\{z_t\}$ as a pulse or gradually increasing (decreasing) process.

iv) non-stationary and you suspect that the intervention has a temporary effect on the level of the $\{y_t\}$.
Answer: Model $\{y_t\}$ as a unit root process and model the intervention variable $\{z_t\}$ such that the values of z_i sum to zero.

v) non-stationary and you suspect that the intervention increases the trend growth of $\{y_t\}$.
Answer: Model $\{y_t\}$ as a unit root process and model the intervention variable $\{z_t\}$ as a pure jump or gradually increasing process.

2. Let the realized value of the $\{z_t\}$ sequence be such that $z_1 = 1$ and all other values of $z_i = 0$.

A. Use equation (5.11) to trace out the effects of the $\{z_t\}$ sequence on the time path of y_t.
Answer: From (2.9) $y_t = a_1y_{t-1} + c_0z_t + \varepsilon_t$ so that:

$$
\begin{aligned}
y_t &= (c_0z_t + \varepsilon_t)/(1\text{-}a_1L) \\
&= c_0(1+a_1L+a_1^2L^2+\ldots+a_1^iL^i+\ldots)\, z_t + \varepsilon_t/(1\text{-}a_1L) \\
&= c_0z_t + c_0a_1z_{t-1} + c_0a_1^2z_{t-2} + \ldots + c_0a_1^iz_{t-i} + \ldots + \varepsilon_t/(1 - a_1L)
\end{aligned}
$$

Since $z_1 = 1$ and all other values of $z_i = 0$:

$$dy_1/dz_1 = c_0,\ \ dy_2/dz_1 = c_0a_1,\ \ dy_3/dz_1 = c_0a_1^2, \text{ and in general: } \boxed{dy_{1+i}/dz_1 = c_0a_1^i}$$

B. Use equation (5.12) to trace out the effects of the $\{z_t\}$ sequence on the time paths of y_t and Δy_t.

Answer: Recall that: $\Delta y_t = a_1 \cdot \Delta y_{t-1} + c_0 z_t + \varepsilon_t$ To trace out the effects of the $\{z_t\}$ sequence on the time path of Δy_t, let $w_t \equiv \Delta y_t$, then write (2.10) as: $w_t = a_1 w_{t-1} + c_0 z_t + \varepsilon_t$.

Since the form of the equation is identical to that in part A, it follows that:

$d(\Delta y_1)/dz_1 = c_0$, $d(\Delta y_2)/dz_1 = c_0 a_1$, $d(\Delta y_3)/dz_1 = c_0 a_1^2$, ..., and in general:

$$\boxed{d(\Delta y_{1+i})/dz_1 = c_0 a_1^i}.$$

To find the effects of the $\{z_t\}$ sequence on the time path of y_t, write (2.10) as follows:

$$y_t - y_{t-1} = a_1(y_{t-1} - y_{t-2}) + c_0 z_t + \varepsilon_t, \text{ or}$$
$$y_t - (1+a_1)y_{t-1} + a_1 y_{t-2} = c_0 z_t + \varepsilon_t, \text{ or}$$
$$[1-(1+a_1)L + a_1L^2]y_t = c_0 z_t + \varepsilon_t.$$

The improper form of the solution for y_t is:

$$\begin{aligned}
y_t &= c_0 z_t/[1-(1+a_1)L+a_1L^2] + \varepsilon_t/[1-(1+a_1)L+a_1L^2] \\
&= c_0 z_t/[(1-L)(1-a_1L)] + \varepsilon_t/[1-(1+a_1)L+a_1L^2] \\
&= c_0[1+(1+a_1)L + (1+a_1+a_1^2)L^2 + ... + (1+a_1+a_1^2+...+a_1^i)L^i + ...]\, z_t \\
&\quad + \varepsilon_t/[1-(1+a_1)L+a_1L^2] \\
&= c_0 z_t + c_0(1+a_1)z_{t-1} + c_0(1+a_1+a_1^2)z_{t-2} + ... + c_0(1+a_1+a_1^2+...+a_1^i)z_{t-i} + ... \\
&\quad + \varepsilon_t/[1-(1+a_1)L+a_1L^2]
\end{aligned}$$

Hence, $dy_t/dz_t = c_0$, $dy_t/dz_{t-1} = dy_{t+1}/dz_t = c_0(1+a_1)$, $dy_t/dz_{t-2} = dy_{t+2}/dz_t = c_0(1+a_1+a_1^2)$, and $dy_t/dz_{t-i} = dy_{t+i}/dz_t = c_0(1+a_1+a_1^2+...+a_1^i)$, ...
It follows that $dy_1/dz_1 = c_0$, $dy_2/dz_1 = c_0(1+a_1)$, $dy_3/dz_1 = c_0(1+a_1+a_1^2)$, ... and:

$$\boxed{dy_{1+i}/dz_1 = c_0(1 + a_1 + a_1^2 + ... + a_1^i).}$$

C. Use (5.13) to trace out the effects of the $\{z_t\}$ sequence on the time patterns of y_t and Δy_t.
Answer: Write (2.11) as follows:

$$y_t - y_{t-1} = a_1(y_{t-1} - y_{t-2}) + c_0(z_t - z_{t-1}) + \varepsilon_t, \text{ or}$$
$$y_t - (1+a_1)y_{t-1} + a_1 y_{t-2} = c_0(z_t - z_{t-1}) + \varepsilon_t, \text{ or}$$
$$[1-(1+a_1)L+a_1L^2]y_t = c_0 \cdot (1-L)z_t + \varepsilon_t.$$

Solving for y_t:

$$\begin{aligned}
y_t &= c_0(1-L)z_t/[1-(1+a_1)L+a_1L^2] + \varepsilon_t/[1-(1+a_1)L+a_1L^2] \\
&= c_0(1-L)z_t/[(1-L)(1-a_1L)] + \varepsilon_t/[(1-L)(1-a_1L)] \\
&= c_0 z_t/(1-a_1L) + \varepsilon_t/[(1-L)(1-a_1L)] \\
&= c_0[1+a_1L+a_1^2L^2+...+a_1^iL^i+...]z_t + \varepsilon_t/[(1-L)(1-a_1L)] \\
&= c_0 z_t + c_0 a_1 z_{t-1} + c_0 a_1^2 z_{t-2} + ... + c_0 a_1^i z_{t-i} + ... + \varepsilon_t/[(1-L)(1-a_1L)]
\end{aligned}$$

Hence, $dy_t/dz_{t-i} = dy_{t+i}/dz_t = c_0a_1^i$. Given that $z_1 = 1$ and all other $z_i = 0$:

$dy_1/dz_1 = c_0$, $dy_2/dz_1 = c_0a_1$, $dy_3/dz_1 = c_0a_1^2$, ...

Next, to find the effects of the $\{z_t\}$ sequence on the time path of Δy_t, rewrite (2.11) as: $\Delta y_t = a_1\Delta y_{t-1} + c_0(1-L)z_t + \varepsilon_t$.

Again, let $\Delta y_t \equiv w_t$ so that:

$(1-a_1L)w_t = c_0(1-L)z_t + \varepsilon_t$.

Solving for w_t (i.e., Δy_t):

$$
\begin{aligned}
w_t &= c_0(1-L)z_t/(1-a_1L) + \varepsilon_t/(1-a_1L) \\
&= c_0[1+(a_1-1)L+a_1(a_1-1)L^2+a_1^2(a_1-1)L^3+...+a^{i-1}(a_1-1)L^i+...]z_t + \varepsilon_t/(1-a_1L) \\
&= c_0z_t+c_0(a_1-1)z_{t-1}+c_0a_1(a_1-1)z_{t-2}+c_0a_1^2(a_1-1)z_{t-3}+... + c_0a_1^{i-1}(a_1-1)z_{t-i}+...+ \varepsilon_t/(1-a_1L).
\end{aligned}
$$

Hence, $dw_t/dz_t = c_0$ and for $i > 0$, $dw_t/dz_{t-i} = dw_{t+i}/dz_t = d\Delta y_{t+i}/dz_t = c_0 \cdot a_1^{i-1}(a_1-1)$. Thus: $d\Delta y_{t+i}/dz_t = c_0\, a_1^{i-1}(a_1-1)$.

Given that $z_1 = 1$ and all other values of $z_i = 0$, it follows that:

$d\Delta y_1/dz_1 = c_0$, $d\Delta y_2/dz_1 = c_0(a_1-1)$, $d\Delta y_3/dz_1 = c_0a_1(a_1-1)$,..., and:

$$\boxed{d\Delta y_{1+i}/dz_1 = c_0a_1^{i-1}(a_1-1).}$$

D. Would your answers to parts A through C change if $\{z_t\}$ was assumed to be a white noise process and you were asked to trace out the effects if a z_t shock of the various $\{y_t\}$ sequences?

Answer: The answers will not change if $\{z_t\}$ was assumed to be a white noise process. Since if we were interested to trace out the effects of a z_t shock (which equals to unity), we only need to change the time index from 1 to t in parts A through C.

E. Assume that $\{z_t\}$ is a white noise process with a variance equal to unity.

(i) Use (5.11) to derive the cross-correlogram between $\{z_t\}$ and $\{y_t\}$.

Answer: From part A, we obtain:

$$y_t = c_0z_t + c_0a_1z_{t-1} + c_0a_1^2z_{t-2} +...+ c_0a_1^iz_{t-i} +...+ \varepsilon_t/(1-a_1L).$$

Form the Yule-Walker equations by multiplying y_t by each z_{t-i}:

$$
\begin{aligned}
&y_tz_t = c_0z_t^2 + c_0a_1z_{t-1}z_t + c_0a_1^2z_{t-2}z_t +...+ c_0a_1^iz_{t-i}z_t +...+ \varepsilon_tz_t/(1-a_1L) \\
&y_tz_{t-1} = c_0z_t \cdot z_{t-1} + c_0a_1z_{t-1}^2 + c_0a_1^2z_{t-2}z_{t-1} +...+ c_0a_1^iz_{t-i}z_{t-1} +...+ \varepsilon_tz_{t-1}/(1-a_1L) \\
&... \\
&y_tz_{t-i} = c_0z_tz_{t-i} + c_0a_1z_{t-1}z_{t-i} + c_0a_1^2z_{t-2}z_{t-i} +...+ c_0a_1^iz_{t-i}^2 +...+ \varepsilon_tz_{t-i}/(1-a_1L\ )
\end{aligned}
$$

...

Since $\{z_t\}$ and $\{\varepsilon_t\}$ are independent white noise disturbances, it follows that:

$Ey_tz_t = c_0$
$Ey_tz_{t-1} = c_0a_1$
...
$Ey_tz_{t-i} = c_0a_1^i$.

The cross-correlation between z_{t-i} and y_t is:

$$\boxed{\begin{array}{l}\rho_{yz}(i) = \text{cov}(y_t, z_{t-i})/(\sigma_y\sigma_z) = c_0a_1^i/\sigma_y\cdot 1 = c_0a_1^i/\sigma_y\ ,\\ where\text{: } \sigma_y = [(c_0^2 + \sigma^2)/(1-a_1^2)]^{1/2},\ \sigma^2 = \text{var}(\varepsilon_t).\end{array}}$$

ii) Use (5.12) to derive the cross-correlogram between $\{z_t\}$ and $\{\Delta y_t\}$.
Answer: Again, let $w_t \equiv \Delta y_t$, so that:

$w_t = a_1w_{t-1} + c_0z_t + \varepsilon_t$, or
$w_t = (c_0z_t + \varepsilon_t)/(1-a_1L)$
$\quad = c_0z_t + c_0a_1z_{t-1} + c_0a_1^2z_{t-2} + ... + c_0a_1^iz_{t-i} + ... + \varepsilon_t/(1-a_1L)$

Form the Yule-Walker by multiplying y_t by each z_{t-i} and take expectations. Since $\{z_t\}$ and $\{\varepsilon_t\}$ are independent:

$Ew_tz_t = c_0$, $Ew_tz_{t-1} = c_0a_1$, and in general, $Ew_tz_{t-i} = c_0a_1^i$.

Hence, the cross-correlogram between $\{z_t\}$ and $\{\Delta y_t\}$ is given by:

$\rho_{\Delta yz}(i) \equiv \text{cov}\ (\Delta y_t, z_{t-i})/(\sigma_{\Delta y}\cdot\sigma_z) = c_0a_1^i/(\sigma_{\Delta y}\cdot 1) = c_0a_1^i/\sigma_{\Delta y}$,
where: $\sigma_{\Delta y} = \sigma_w = [(c_0^2 + \sigma^2)/(1-a_1^2)]^{1/2}$

Note that the pattern of the cross-correlogram is the same as that above, i.e.,

$$\boxed{\rho_{\Delta yz}(i) \equiv c_0(a_1)^i/[(c_0^2 + \sigma^2)/(1-a_1^2)]^{1/2}}$$

(iii) Use (5.13) to derive the cross-correlogram between $\{z_t\}$ and $\{\Delta y_t\}$.
Answer: Let $w_t \equiv \Delta y_t$, and using lag operators to write: $(1-a_1L)w_t = c_0(1-L)z_t + \varepsilon_t$.

In part C it was shown that:

$w_t = c_0z_t + c_0(a_1-1)z_{t-1} + c_0a_1(a_1-1)z_{t-2} + ... + c_0a_1^{i-1}(a_1-1)z_{t-i} + ... + \varepsilon_t/(1-a_1L)$.

Forming the Yule-Walker equations and noting that $\{z_t\}$ and $\{\varepsilon_t\}$ are independent white noise disturbances, it follows that:

$Ew_tz_t = E\Delta y_tz_t = c_0$
$Ew_tz_{t-1} = E\Delta y_tz_{t-1} = c_0(a_1\text{-}1)$
$Ew_tz_{t-2} = E\Delta y_tz_{t-2} = c_0a_1(a_1\text{-}1)$
...
$Ew_tz_{t-i} = E\Delta y_tz_{t-i} = c_0a_1^{i\text{-}1}(a_1\text{-}1)$.

Hence, for $i = 0$, $\rho_{\Delta yz}(0) = c_0/\sigma_{\Delta y}$ and for $i > 0$, the cross-correlations are:

$$\boxed{\begin{array}{l}\rho_{\Delta yz}(i) \equiv \text{cov}(\Delta y_t, z_{t-i})/(\sigma_{\Delta y}\cdot\sigma_z) = c_0a_1^{i\text{-}1}(a_1\text{-}1)/\sigma_{\Delta y}\text{ , for } i \geq 1\\ where\text{: } \sigma_{\Delta y} = \sigma_W = [2c_0^2/(1+a_1) + \sigma^2/(1\text{-}a_1^2)]^{1/2}\end{array}}$$

(iv) Now suppose that z_t is the random walk process $z_t = z_{t-1} + \varepsilon_{zt}$. Trace out the effects of an ε_{zt} shock on the Δy_t sequence.
Answer: Case (1): $\Delta y_t = a_1\Delta y_{t-1} + c_0z_t + \varepsilon_t$. Since $(1\text{-}L)z_t = \varepsilon_{zt}$, it follows that:

$\Delta y_t = a_1\Delta y_{t-1} + c_0\varepsilon_{zt}/(1\text{-}L) + \varepsilon_t$ or letting $w_t = \Delta y_t$:
$(1\text{-}a_1L)w_t = c_0\varepsilon_{zt}/(1\text{-}L) + \varepsilon_t$. Solve for w_t to obtain:
$w_t = c_0\varepsilon_{zt}/[(1\text{-}L)(1\text{-}a_1L)] + \varepsilon_t/(1\text{-}a_1L)$

The form of the impulse responses is identical to that in part B. Hence:

$$\boxed{d\Delta y_{t+i}/d\varepsilon_{zt} = c_0\sum_{j=0}^{i} a_1^j}$$

Case (2): $\Delta y_t = a_1\Delta y_{t-1} + c_0\Delta z_t + \varepsilon_t$. Define $w_t = \Delta y_t$ and note that $\Delta z_t = \varepsilon_{zt}$. Hence:

$w_t = a_1w_{t-1} + c_0\varepsilon_{zt} + \varepsilon_t$.

Notice the same pattern as in the part A, so that:

$dw_t/d\varepsilon_{zt} = d\Delta y_t/d\varepsilon_{zt} = c_0$ and for $i > 0$:

$$\boxed{d\Delta y_{t+i}/d\varepsilon_{zt} = c_0a_1^i}$$

3. Consider the transfer function model: $y_t = 0.5\, y_{t-1} + z_t + \varepsilon_t$ where z_t is the autoregressive process: $z_t = 0.5\, z_{t-1} + \varepsilon_{zt}$

A. Derive the cross-correlation between the filtered $\{y_t\}$ sequence and the $\{\varepsilon_{zt}\}$ sequence.

Answer: Since $z_t = 0.5\, z_{t-1} + \varepsilon_{zt}$, we may write the $\{z_t\}$ sequence as: $(1\text{-}0.5L)z_t = \varepsilon_{zt}$. In terms of the notation used in the text, $D(L) \equiv (1\text{-}0.5L)$ and $E(L) \equiv 1$, so that $D(L)z_t = E(L)\varepsilon_{zt}$. Also let $A(L) = 0.5$, $C(L) = 1$, and $B(L) = 1$, so that:

$y_t = A(L)y_{t-1} + C(L)z_t + B(L)\varepsilon_t$. Here, the filtered value of y_t (y_{ft}) is:

$y_{ft} = [D(L)/E(L)]y_t = (1\text{-}0.5L)y_t$
$= 0.5 \cdot (1\text{-}0.5L)y_{t-1} + (1\text{-}0.5L)z_t + (1\text{-}0.5L)\varepsilon_t$, so that:
$y_{ft} = 0.5y_{ft\text{-}1} + \varepsilon_{zt} + (1\text{-}0.5L)\varepsilon_t$.

A simple way to obtain the cross-covariances is to multiply y_{ft} by the successive values of ε_{zt-i}. Since $E\varepsilon_t\varepsilon_{zt-i} = 0$ for all t and i, it follows that:

$Ey_{ft}\varepsilon_{zt} = \sigma^2_{\varepsilon z}$
$Ey_{ft}\varepsilon_{zt\text{-}1} = 0.5\sigma^2_{\varepsilon z}$
$Ey_{ft}\varepsilon_{zt\text{-}2} = 0.5^2\ \sigma^2_{\varepsilon z}$
$Ey_{ft}\varepsilon_{zt\text{-}3} = 0.5^3\ \sigma^2_{\varepsilon z}$, so that in compact form:
$Ey_{ft}\varepsilon_{zt\text{-}i} = 0.5^i\ \sigma^2_{\varepsilon z}$.

The cross-correlation between y_{ft} and ε_{zt-i} is defined as: $\rho_{yz}(i) = \mathrm{Cov}(y_{ft}, \varepsilon_{zt\text{-}i})/(\sigma_{yf}\sigma_{\varepsilon z})$. Since $Ey_{ft} = 0$, $E\varepsilon_{zt} = 0$, and $\mathrm{var}(y_{ft}) = (1{+}0.5^2{+}0.5^4{+}0.5^6{+}...)\sigma^2_{\varepsilon z} + \sigma^2$, it follows that $\sigma_{yf} = (4/3\ \sigma^2_{\varepsilon z} + \sigma^2)^{1/2}$. Hence, the cross-correlations between the filtered $\{y_t\}$ sequence and the $\{\varepsilon_{zt}\}$ sequence are:

$\rho_{yz}(i) = 0.5^i\ \sigma^2_{\varepsilon z}/[(4/3\ \sigma^2_{\varepsilon z} + \sigma^2)^{1/2} \cdot \sigma_{\varepsilon z}]$ or:
$$\boxed{\rho_{yz}(i) = 0.5^i \cdot \sigma_{\varepsilon z}/(4/3\ \sigma^2_{\varepsilon z} + \sigma^2)^{1/2}}$$

B. Now suppose $y_t = 0.5\ y_{t\text{-}1} + z_t + 0.5\ z_{t\text{-}1} + \varepsilon_t$ and $z_t = 0.5\ z_{t\text{-}1} + \varepsilon_{zt}$ .Derive the cross-autocovariances between the filtered $\{y_t\}$ sequence and ε_{zt}. Show that the first two cross-autocovariances are proportional to the transfer function coefficients. Show that the cross-covariances decay at the rate 0.5.

Answer: Since $z_t = 0.5\ z_{t\text{-}1} + \varepsilon_{zt}$, we can write $(1\text{-}0.5L)\ z_t = \varepsilon_{zt}$. The filtered values of y_t sequence (i.e., y_{ft}) are: $y_{ft} = (1\text{-}0.5L)\ y_t$. Hence:

$y_{ft} = 0.5y_{ft\text{-}1} + (1 + 0.5L)\varepsilon_{zt} + (1 - 0.5L)\ \varepsilon_t$.

You can obtain the Yule-Walker equations in precisely the same way as in part A above. For an alternative solution technique, obtain the moving average representation for y_{ft} as:

$(1\text{-}0.5L)\ y_{ft} = (1{+}0.5L)\ \varepsilon_{zt} + (1\text{-}0.5L)\ \varepsilon_t$ so that:

$y_{ft} = (1{+}0.5L)\varepsilon_{zt}/(1\text{-}0.5L) + \varepsilon_t$
$= (1{+}L + 0.5L^2 + 0.5^2L^3 + 0.5^3L^4 + 0.5^4L^5 +...)\ \varepsilon_{zt} + \varepsilon_t$
$= \varepsilon_{zt} + \varepsilon_{zt\text{-}1} + 0.5\ \varepsilon_{zt\text{-}2} + 0.5^2\ \varepsilon_{zt\text{-}3} + 0.5^3\ \varepsilon_{zt\text{-}4} + 0.5^4\ \varepsilon_{zt\text{-}5} +... + \varepsilon_t$.

Now multiply by the successive values of ε_{zt-i} to obtain:

$y_{ft}\varepsilon_{zt} = \varepsilon_{zt}\cdot\varepsilon_{zt} + \varepsilon_{zt\text{-}1}\cdot\varepsilon_{zt} + 0.5\varepsilon_{zt\text{-}2}\cdot\varepsilon_{zt} + 0.5^2\varepsilon_{zt\text{-}3}\cdot\varepsilon_{zt} +...+ \varepsilon_t\cdot\varepsilon_{zt}$

$y_{ft}\varepsilon_{zt\text{-}1} = \varepsilon_{zt}\cdot\varepsilon_{zt\text{-}1} + \varepsilon_{zt\text{-}1}\cdot\varepsilon_{zt\text{-}1} + 0.5\varepsilon_{zt\text{-}2}\cdot\varepsilon_{zt\text{-}1} + 0.5^2\cdot\varepsilon_{zt\text{-}3}\cdot\varepsilon_{zt\text{-}1} + 0.5^3\ \varepsilon_{zt\text{-}4}\cdot\varepsilon_{zt\text{-}1} + ... + \varepsilon_t\cdot\varepsilon_{zt\text{-}1}$
$y_{ft}\varepsilon_{zt\text{-}2} = \varepsilon_{zt}\cdot\varepsilon_{zt\text{-}2} + \varepsilon_{zt\text{-}1}\cdot\varepsilon_{zt\text{-}2} + 0.5\ \varepsilon_{zt\text{-}2}\cdot\varepsilon_{zt\text{-}2} + 0.5^2\cdot\varepsilon_{zt\text{-}3}\cdot\varepsilon_{zt\text{-}2} + 0.5^3\ \varepsilon_{zt\text{-}4}\cdot\varepsilon_{zt\text{-}2} + ... + \varepsilon_t\cdot\varepsilon_{zt\text{-}2}$
$y_{ft}\ \varepsilon_{zt\text{-}3} = \varepsilon_{zt}\cdot\varepsilon_{zt\text{-}3} + \varepsilon_{zt\text{-}1}\cdot\varepsilon_{zt\text{-}3} + 0.5\ \varepsilon_{zt\text{-}2}\cdot\varepsilon_{zt\text{-}3} + 0.5^2\cdot\varepsilon_{zt\text{-}3}\cdot\varepsilon_{zt\text{-}3} + 0.5^3\ \varepsilon_{zt\text{-}4}\cdot\varepsilon_{zt\text{-}3} + ... + \varepsilon_t\cdot\varepsilon_{zt\text{-}3}$
...

Now take the expected value of each of the above equations and note that $\{\varepsilon_t\}$ and $\{\varepsilon_{zt}\}$ are independent white noise disturbances. It follows that:

$Ey_{ft}\varepsilon_{zt} = \sigma^2_{\varepsilon z}$,
$Ey_{ft}\varepsilon_{zt\text{-}1} = \sigma^2_{\varepsilon z}$,
$Ey_{ft}\varepsilon_{zt\text{-}2} = 0.5\ \sigma^2_{\varepsilon z}$,
$Ey_{ft}\varepsilon_{zt\text{-}3} = 0.5^2\ \sigma^2_{\varepsilon z}$
...

$$\boxed{Ey_{ft}\varepsilon_{zt\text{-}i}\ =\ 0.5^{i\text{-}1}\ \sigma^2_{\varepsilon z}\ \text{ for all } i \geq 2}$$

Clearly, the cross-autocovariances decay at the rate of 0.5. Since $c_0 = 1$, it is obvious that the first two autocovariances are proportional to the transfer function coefficients.

4. Use (5.28) to find the appropriate second-order stochastic difference equation for y_t.

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} y_{t1} \\ z_{t1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}$$

One way to answer the questions is to first transform the first-order VAR in the $\{y_t\}$ and $\{z_t\}$ sequences into a second-order difference equation for y_t. Using (5.10), it follows that:

$y_t =\ [(1\text{-}0.8L)e_{1t} + 0.2e_{2t\text{-}1}]\ /[(1\text{-}0.8L)(1\text{-}0.8L) - 0.04\ L^2]$.

Similarly, the solution for z_t is:

$z_t\ =\ [0.2\ e_{1\ t\text{-}1} + (1\text{-}0.8L)e_{2t}]\ /[(1\text{-}0.8L)(1\text{-}0.8L) - 0.04\ L^2]$.

A. Determine whether the $\{y_t\}$ sequence is stationary.
Answer: The inverse characteristic equation is: $(1\text{-}0.8L)(1\text{-}0.8L) - 0.04\ L^2 = 0$. The two roots of the inverse characteristic equation are 1 and 5/3. Although one root (5/3) is outside the unit circle, the other root (1) is on the unit circle, the sequence is not stationary.

B. Discuss the shape of the impulse response function of y_t to a one unit shock in e_{1t} and to a one unit shock in e_{2t}.
Answer: Iterating backwards from t yields:
$y_t = (1{+}0.8L + 0.8\cdot0.6L^2 + 0.8\cdot0.6^2L^3 + 0.8\cdot0.6^3L^4 + ...)\ e_{1t}$
$+ (0.2L + 0.32L^2 + 0.392L^3 + 0.4352L^4 + 0.46112L^5 + ...)e_{2t}$.
Hence,
$dy_{t+1}/de_{1t} = dy_t/de_{1t\text{-}1} = 0.8$; $\qquad dy_{t+1}/de_{2t} = dy_t/de_{2t\text{-}1} = 0.2$

$dy_{t+2}/de_{1t} = dy_t/de_{1t-2} = 0.8 \cdot 0.6 = 0.48$; $dy_{t+2}/de_{2t} = dy_t/de_{2t-2} = 0.32$
$dy_{t+3}/de_{1t} = dy_t/de_{1t-3} = 0.8 \cdot 0.6^2 = 0.288$; $dy_{t+3}/de_{2t} = dy_t/de_{2t-3} = 0.392$
$dy_{t+4}/de_{1t} = dy_t/de_{1t-4} = 0.8 \cdot 0.6^3 = 0.1728$; $dy_{t+4}/de_{2t} = dy_t/de_{2t-4} = 0.4352$
...

C. Suppose $e_{1t} = \varepsilon_{yt} + 0.5\,\varepsilon_{zt}$ and $e_{2t} = \varepsilon_{zt}$. Discuss the shape of the impulse response function of y_t to a one-unit shock in ε_{yt}. Repeat for a one-unit shock in ε_{zt}.

Answer: The responses to ε_{yt} shocks are identical to those given in part B above. However, in response to a one-unit shock in ε_{zt}, ε_{yt} changes by 0.5 units. Hence, the response to an ε_{zt} shock is equal to that reported in part B plus 0.5 times the effect of an ε_{yt} shock on y_{t+i}, i.e.,

$$dy_{t+1}/d\varepsilon_{zt} = (0.8/2 + 0.2) = 0.6,\ dy_{t+2}/d\varepsilon_{zt} = (0.48/2 + 0.32) = 0.56,$$
$$dy_{t+3}/d\varepsilon_{zt} = (0.288/2 + 0.392) = 0.536,$$
$$dy_{t+4}/d\varepsilon_{zt} = (0.1728/2 + 0.4352) = 0.5216$$

D. Suppose $e_{1t} = \varepsilon_{yt}$ and that $e_{2t} = 0.5\varepsilon_{yt} + \varepsilon_{zt}$. Discuss the shape of the impulse response function of y_t to a one-unit shock in ε_{yt}. Repeat for a one-unit shock in ε_{zt}.

Answer: The responses to an ε_{zt} shock are identical to those given in part B above. However, in response to a one-unit shock in ε_{yt}, ε_{zt} changes by 0.5 units. Hence, the response to an ε_{yt} shock is equal to that reported in part B plus 0.5 times the effect of an ε_{zt} shock on y_{t+i}, i.e.,

$$dy_{t+1}/d\varepsilon_{yt} = (0.8 + 0.2/2) = 0.9,\ dy_{t+2}/d\varepsilon_{yt} = (0.48 + 0.32/2) = 0.64$$
$$dy_{t+3}/d\varepsilon_{yt} = (0.288 + 0.392/2) = 0.484,$$
$$dy_{t+4}/d\varepsilon_{yt} = (0.1728 + 0.4352/2) = 0.3904$$

E. Use your answers to C and D to explain why the ordering in Choleski decomposition is important.

Answer: In essence, part C uses a Choleski decomposition such that the current value of y_t does not have a contemporaneous effect on z_t. In part D the current value of z_t does not have a contemporaneous effect on y_t. Both models show that the effect of a one unit shock of ε_{yt} will eventually die out (with different rates). However, the impulse responses of $\{y_t\}$ to an ε_{zt} shock are profoundly different. In C, the effects of ε_{zt} shocks eventually die out; in parts B and D, the effects of ε_{zt} shocks increase without bound.

F. Using the notation in (5.27), find $(A_1)^2$ and $(A_1)^3$. Does $(A_1)^n$ appear to approach zero (i.e., the null matrix)?

Answer:

$$A_1 A_1 = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.68 & 0.32 \\ 0.32 & 0.68 \end{bmatrix}$$

and:

$$A_1^3 = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.68 & 0.32 \\ 0.32 & 0.68 \end{bmatrix} = \begin{bmatrix} 0.608 & 0.392 \\ 0.392 & 0.608 \end{bmatrix}$$

The expression $(A_1)^n$ does not approach zero as n increases. Instead $(A_1)^n$ approaches:

$$A_1^n = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

Thus, for large n, $A_1(A_1)^n = (A_1)^n$. To verify this result, note that:

$$\begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

5. Using the notation of (5.20) and (5.21) suppose: $a_{10} = 0$, $a_{20} = 0$, $a_{11} = 0.8$, $a_{12} = 0.2$, $a_{21} = 0.4$, and $a_{22} = 0.1$.

A. Find the appropriate second-order stochastic difference equation for y_t. Determine whether the $\{y_t\}$ sequence is stationary.

Answer: Using the same technique as in question 4, the solutions for y_t and z_t are:

$$y_t = [(1\text{-}0.1L)e_{1t} + 0.2Le_{2t}]/[(1\text{-}0.8L)(1\text{-}0.1L) - 0.08\,L^2]$$
$$z_t = [0.4Le_{1t} + (1\text{-}0.8L)e_{2t}]/[(1\text{-}0.8L)(1\text{-}0.1L) - 0.08\,L^2]$$

Notice that the inverse characteristic equation has a single root; i.e.,

$$(1\text{-}0.8L)(1\text{-}0.1L) - 0.08\,L^2 = 1 - 0.9L$$

Hence, the solutions for y_t and z_t are:

$$\boxed{\begin{aligned} y_t &= [(1\text{-}0.1L)e_{1t} + 0.2Le_{2t}]/(1 - 0.9L) \\ z_t &= [0.4Le_{1t} + (1\text{-}0.8L)e_{2t}]/(1 - 0.9L) \end{aligned}}$$

Since the root of the inverse characteristic equation is 10/9, the sequence is stationary.

B. Answer the parts B through F of Question 4 using these new values of the a_{ij}

Part B: Discuss the shape of the impulse response function of y_t to a one unit shock in e_{1t} and to a one unit shock in e_{2t}.

Answer: Write the solution for y_t in its moving-average form:

$$y_t = (1{+}0.8L + 0.8\cdot0.9L^2 + 0.8\cdot0.9^2L^3 + 0.8\cdot0.9^3L^4 + ...)\, e_{1t}$$
$$+ (0.2L + 0.2\cdot0.9L^2 + 0.2\cdot0.9^2L^3 + 0.2\cdot0.9^3L^4 + ...)\, e_{2t}\,.$$

Hence the impulse responses are:

$dy_{t+1}/de_{1t} = dy_t/de_{1t-1} = 0.8$, $\quad dy_{t+1}/de_{2t} = dy_t/de_{2t-1} = 0.2$

$dy_{t+2}/de_{1t} = dy_t/de_{1t-2} = 0.8\cdot0.9 = 0.72$, $\quad dy_{t+2}/de_{2t} = dy_t/de_{2t-2} = 0.2\cdot0.9 = 0.18$

$dy_{t+3}/de_{1t} = dy_t/de_{1t-3} = 0.8\cdot0.9^2 = 0.648$, $\quad dy_{t+3}/de_{2t} = dy_t/de_{2t-3} = 0.2\cdot0.9^2 = 0.162$

$dy_{t+4}/de_{1t} = dy_t/de_{1t-4} = 0.8\cdot0.9^3 = 0.5832$, $\quad dy_{t+4}/de_{2t} = dy_t/\sigma e_{2t-4} = 0.2\cdot0.9^3 = 0.1458$

...

Part C. Suppose $e_{1t} = \varepsilon_{yt} + 0.5\varepsilon_{zt}$ and $e_{2t} = \varepsilon_{zt}$. Discuss the shape of the impulse response function of y_t to a one unit shock in ε_{yt}. Repeat for a one unit shock in ε_{zt}.

Answer: The effects of an ε_{yt} are unchanged from the answer in part B. Here, the ε_{zt} shock induces a 0.5 change in e_{1t}. Hence, the response to an ε_{zt} shock is equal to that reported in part B plus 0.5 times the effect of an e_{1t} shock on y_{t+i}:

$$\boxed{\begin{array}{l} dy_{t+1}/d\varepsilon_{zt} = (0.8/2 + 0.2) = 0.6,\ \ dy_{t+2}/d\varepsilon_{zt} = (0.72/2 + 0.18) = 0.54 \\ dy_{t+3}/d\varepsilon_{zt} = (0.648/2 + 0.162) = 0.486 \\ dy_{t+4}/d\varepsilon_{zt} = (0.5832/2 + 0.1458) = 0.4374, ... \end{array}}$$

Part D. Suppose $e_{1t} = \varepsilon_{yt}$ and $e_{2t} = 0.5\varepsilon_{yt} + \varepsilon_{zt}$. Discuss the shape of the impulse response function of y_t to a one unit shock in ε_{yt}. Repeat for a one unit shock in ε_{zt}.

Answer: Now, the effects of an ε_{zt} are unchanged from the answer in part B. However, the ε_{yt} shock induces a 0.5 change in e_{2t}. Hence, the response to an ε_{yt} shock is equal to that reported in part B plus 0.5 times the effect of an e_{2t} shock on y_{t+i}:

$dy_{t+1}/d\varepsilon_{yt} = 0.9$, $dy_{t+2}/d\varepsilon_{yt} = 0.81$, and in general:

$$\boxed{dy_{t+i}/d\varepsilon_{yt} = dy_t/d\varepsilon_{yt-i} = (0.9)^i}$$

Part E. Use your answers to parts C and D to explain why the ordering in a Choleski decomposition is important.

Answer: A Choleski decomposition equates 100% of the 1-step ahead forecast error in one of the variables with a pure innovation in that variable. In part C, for example, e_{2t} is equivalent to an innovation in ε_{zt}. The triangular structure means that an e_{2t} shock has a contemporaneous effect on y_t and z_t but that an e_{1t} shock has a contemporaneous effect only on y_t. Since the VAR innovations cannot be identified, the restriction imposed by Choleski decomposition is not testable. Part D introduces the alternative Choleski decomposition. Obviously, the decay patterns present using the alternative decompositions are quite different.

Part F. Using the notation in (5.19), find $(A_1)^2$ and $(A_1)^3$. Does $(A_1)^n$ appear to approach zero (i.e., the null matrix) ?

Answer:

$$A_1\,A_1 = \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.1 \end{bmatrix}\begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.1 \end{bmatrix} = \begin{bmatrix} 0.72 & 0.18 \\ 0.36 & 0.09 \end{bmatrix}$$

$$A_1^3 = \begin{bmatrix} 0.72 & 0.18 \\ 0.36 & 0.09 \end{bmatrix}\begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.648 & 0.162 \\ 0.324 & 0.081 \end{bmatrix}$$

As opposed to the answer in question 4, here $(A_1)^n$ does approach zero.

C. How would the solution for y_t change if $a_{10} = 0.2$?

Answer: The new solution for y_t is:

$$y_t = [0.18 + (1\text{-}0.1L)e_{1t} + 0.2Le_{2t}] /(1\text{-}0.9L)$$
$$= 1.8 + [(1\text{-}0.1L)\, e_{1t} + 0.2Le_{2t}\,]/(1\text{-}0.9L)$$

Here, y_t has a non-zero mean. However, the impulse responses are invariant to the non-zero intercept.

6. Suppose the residuals of a VAR are such that Var(e_1) = 0.75, Var(e_2) = 0.5 and Cov(e_{1t}, e_{2t}) = 0.25.

A. Using (5.55)-(5.58) as guides, show that it is not possible to identify the structural VAR.

Answer: Given that:

$$\Sigma = \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.5 \end{bmatrix}$$

it follows that:

$$\begin{bmatrix} \text{var}(\varepsilon_1) & 0.0 \\ 0.0 & \text{var}(\varepsilon_2) \end{bmatrix} = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.5 \end{bmatrix} \begin{bmatrix} 1 & b_{21} \\ b_{12} & 1 \end{bmatrix}$$

Carry out the indicated multiplication to obtain:

$$\text{Var}(\varepsilon_1) = 0.75 + 0.5b_{12} + 0.5b^2_{12.}$$
$$0 = 0.25 + 0.5b_{12} + 0.75b_{21} + 0.25b_{12}b_{21.}$$
$$0 = 0.25 + 0.5b_{12} + 0.75b_{21} + 0.25b_{12}b_{21.}$$
$$\text{Var}(\varepsilon_2) = 0.5 + 0.5b_{21} + 0.75b^2_{21.}$$

The second and third are identical. Hence, there are three independent equations to solve for the four unknowns b_{12}, b_{21}, Var(ε_1), and Var(ε_2). Without an additional restriction, it is not possible to identify the structural VAR.

B. Using Choleski decomposition such that $b_{12} = 0$, find the identified values of b_{21}, Var(ε_1), and Var(ε_2).

Answer: If $b_{12} = 0$, the three independent equations in part A yield:

$$\text{Var}(\varepsilon_1) = 0.75$$
$$0 = 0.25 + 0.75b_{21}, \text{ so that: } b_{21} = -1/3.$$
$$\text{Var}(\varepsilon_2) = 0.5 + 0.5b_{21} + 0.75b^2_{21} = 5/12.$$

C. Using Choleski decomposition such that $b_{21} = 0$, find the identified values of b_{12}, Var(ε_1), and Var(ε_2).

Answer: If $b_{21} = 0$, the three independent equations in part A yield:

$\text{Var}(\varepsilon_2) = 0.5.$
$0 = 0.25 + 0.5b_{12}$, so that: $b_{12} = -1/2.$
$\text{Var}(\varepsilon_1) = 0.75 + 0.5b_{12} + 0.5b^2_{12}$, so that $\text{Var}(\varepsilon_1) = 5/8.$

D. Using a Sims-Bernanke decomposition such that $b_{12} = 0.5$, find the identified values of b_{21}, $\text{Var}(\varepsilon_1)$, and $\text{Var}(\varepsilon_2)$.

Answer: If $b_{12} = 0.5$, we find:

$\text{Var}(\varepsilon_1) = 0.75 + 0.5\cdot 0.5 + 0.5\cdot(0.5)^2 = 1.125.$
$0 = 0.25 + 0.5\cdot 0.5 + 0.75b_{21} + 0.25\cdot 0.5b_{21}$, so that: $b_{21} = -6/7.$
$\text{Var}(\varepsilon_2) = 0.5 + 0.5b_{21} + 0.75b^2_{21}$, so that $\text{Var}(\varepsilon_2) = 61/98.$

E. Using a Sims-Bernanke decomposition such that $b_{21} = 0.5$, find the identified values of b_{12}, $\text{Var}(\varepsilon_1)$, and $\text{Var}(\varepsilon_2)$.

Answer: If $b_{21} = 0.5$, we find:

$\text{Var}(\varepsilon_2) = 0.5 + 0.5\cdot 0.5 + 0.75\cdot(0.5)^2 = 0.9375.$
$0 = 0.25 + 0.5b_{12} + 0.75\cdot 0.5 + 0.25\cdot 0.5\cdot b_{12}$, so that: $b_{12} = -1.$
$\text{Var}(\varepsilon_1) = 0.75 + 0.5b_{12} + 0.5b^2_{12}$, so that $\text{Var}(\varepsilon_1) = 0.75 + 0.5\cdot(-1) + 0.5\cdot(-1)^2 = 0.75.$

F. Suppose that the first three values of e_{1t} are estimated to be 1, 0, -1 and that the first three values of e_{2t} are estimated to be -1, 0, 1. Find the first three values of ε_{1t} and ε_{2t} using each of the decomposition in parts B through E.

Answer: Given that the first three errors terms are:

t	e_{1t}	e_{2t}
1	1.0	-1.0
2	0.0	0.0
3	-1.0	1.0

then $\sigma^2_1 = 2/3$, $\sigma_{12} = \sigma_{21} = -2/3$, and $\sigma^2_2 = 2/3$. Hence, the variance/covariance matrix is:

$$\Sigma = \begin{bmatrix} 2/3 & -2/3 \\ -2/3 & 2/3 \end{bmatrix} \text{ so that}$$

$$\begin{bmatrix} \text{var}(\varepsilon_1) & 0.0 \\ 0.0 & \text{var}(\varepsilon_2) \end{bmatrix} = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} 2/3 & -2/3 \\ -2/3 & 2/3 \end{bmatrix} \begin{bmatrix} 1 & b_{21} \\ b_{12} & 1 \end{bmatrix}$$

Carry out the indicated multiplication to obtain:

$\text{Var}(\varepsilon_1) = 2/3 - 4/3\ b_{12} + 2/3\ b^2_{12.}$
$0 = -2/3 + 2/3\ b_{12} + 2/3\ b_{21} - 2/3\ b_{12}b_{21.}$

$0 \;=\; -2/3 + 2/3\, b_{12} + 2/3\, b_{21} - 2/3\, b_{12}b_{21}.$
$\text{Var}(\varepsilon_2) \;=\; 2/3 - 4/3\, b_{21} + 2/3\, b^2_{21}.$

If $b_{12} = 0$, we find:

$\text{Var}(\varepsilon_1) \;=\; 2/3.$
$0 \;=\; -2/3 + 2/3\, b_{21}$, so that: $b_{21} = 1$.
$\text{Var}(\varepsilon_2) \;=\; 2/3 - 4/3\, b_{21} + 2/3\, b^2_{21}$ so that $\text{Var}(\varepsilon_2) = 2/3 - 4/3 + 2/3 = 0$.

Using the notation of the text, each $\{\varepsilon_{1t}\}$ and $\{\varepsilon_{2t}\}$ sequence can be recovered as $\varepsilon_t = Be_t$ so that:

$\varepsilon_{1t} \;=\; e_{1t}$ and
$\varepsilon_{2t} \;=\; e_{1t} + e_{2t}$.

Thus, the identified structural shocks are:

t	ε_{1t}	ε_{2t}
1	1.0	0.0
2	0.0	0.0
3	-1.0	0.0

If $b_{21} = 0$, we find

$\text{Var}(\varepsilon_2) \;=\; 2/3.$
$0 \;=\; -2/3 + 2/3\, b_{12}$, so that: $b_{12} = 1$.
$\text{Var}(\varepsilon_1) \;=\; 2/3 - 4/3\, b_{12} + 2/3\, b^2_{12}$, so that $\text{Var}(\varepsilon_1) = 0$.

Using the notation of the text, each $\{\varepsilon_{1t}\}$ and $\{\varepsilon_{2t}\}$ sequence can be recovered as $\varepsilon_t = Be_t$ so that:

$\varepsilon_{1t} \;=\; e_{1t} + e_{2t}$ and:
$\varepsilon_{2t} \;=\; e_{2t}$.

Thus, the identified structural shocks are:

t	ε_{1t}	ε_{2t}
1	0.0	-1.0
2	0.0	0.0
3	0.0	1.0

If $b_{12} = 0.5$, you can use the same logic to find:

t	ε_{1t}	ε_{2t}
1	0.5	0.0
2	0.0	0.0
3	-0.5	0.0

If $b_{21} = 0.5$, you should find:

t	ε_{1t}	ε_{2t}
1	0.0	-0.5
2	0.0	0.0
3	0.0	0.5

7. This set of exercises uses data from the file entitled QUARTELY.XLS. In Chapter 2 the logarithmic changes in the U.S. Producer Price Index and M1 (not seasonally adjusted) were both estimated using the Box-Jenkins method. Here, the goal is to model both simultaneously using a VAR.

Program for RATS Users

```
* The first four lines read in the data set

cal 1960 1 4
all 4 2002:1
open data a:\quarterly.xls
data(format=xls,org=obs)

*  Next, create the growth rate of the money supply (gm1) and the inflation rate (inf)
set gm1 = log(m1nsa) - log(m1nsa{1})
set inf = log(ppi) - log(ppi{1})
seasonal seasons                         ; * Create the seasonal dummy variable

* First estimate the VAR with 12 lags
system 1 to 2
vars gm1 inf
lags 1 to 12
det constant seasons{-2 to 0}
end(system)
estimate(outsigma=v,noprint,sigma) / 1

* Now use 8 lags. Be sure to begin the estimation in 1963Q2 so that the two systems are
* estimated over the same sample period.
system 1 to 2
vars gm1 inf
lags 1 to 8
det constant seasons{-2 to 0}
```

```
end(system)
estimate(noprint) / 3
* Perform the likelihood ratio test
ratio(degrees=16,mcorr=28) 63:2 *          ;* There are 16 degrees of freedom in that 4 lags of each
# 1 2                                      ;* variable are eliminated from each equation. Each
# 3 4                                      ;*equation in the 12-lag system has 28 regressors

* The test for 8 to 4 lags can be conducted using
system 1 to 2
vars gm1 inf
lags 1 to 8
det constant seasons{-2 to 0}
end(system)
estimate(noprint) / 1

system 1 to 2
vars gm1 inf
lags 1 to 4
det constant seasons{-2 to 0}
end(system)
estimate(noprint) 62:2 * 3
ratio(degrees=16,mcorr=20) 62:2 *          ;* Estimate beginning with 1962Q2. Now the small
# 1 2                                      ;* correction is 20 since the 8-lag equations contain 20
# 3 4                                      ;* regressors (8 lags of each variable plus the intercept
                                           ;* plus the three seasonals

* The 8-lag model including the seasonals and intercepts can be estimated using
system 1 to 2
vars gm1 inf
lags 1 to 8
det constant seasons{-2 to 0}
end(system)
estimate(outsigma=V,sigma) / 1
* The following can be used to produce the impulse responses and variance decompositions
errors(impulses) 2 24 V
# 1 ; # 2

*** To estimate the VAR in levels use
set m1 = log(m1nsa)
seasonal seasons
set lp = log(ppi)

system 1 to 2
```

```
 vars m1 lp
 lags 1 to 12
 det constant seasons{-2 to 0}
 end(system)
 estimate / 1

 errors(impulses) 2 24 V
 # 1 ; # 2
```

Program for STATA Users

```
clear
cd "x:\New_data"  /*change this line if the dataset is saved elsewhere*/
insheet using quarterly.txt
gen obs = _n
tsset obs /*declare this is a time series dataset*/
gen lm1nsa = ln(m1nsa)
gen lppi = ln(ppi)
gen mt = d.lm1nsa /*this is the rate of growth of the money supply*/
gen piet = d.lppi /*this is the inflation rate*/

/*generate three seasonal dummy variables*/
gen d1=0
quietly replace d1= 1 if match(descriptor,"*Q1")
gen d2=0
quietly replace d2= 1 if match(descriptor,"*Q2")
gen d3=0
quietly replace d3= 1 if match(descriptor,"*Q3")

/*B:estimate the VARs*/
set matsize 800
quietly var mt piet,lags(1/12) exog(d1 d2 d3)
scalar ldet12 = ln(e(detsig))
display ldet12
scalar numpar12 = e(tparms)/e(neqs)
quietly var mt piet in 14/169,lags(1/8) exog(d1 d2 d3)
scalar ldet8 = ln(e(detsig))
display ldet8  /*ii:estimate the VARs over the same period (1963Q2 - 2002Q1)using 12 and 8 lags respectively*/

/*iii:perform the likelihood ratio test for the null hypothesis of 8 lags*/
scalar lr8 = (e(T)-numpar12)*(ldet8-ldet12)
display "likelihood ratio = "= lr8
```

```
display "the significance level of the likelihood ratio stat = "= chi2tail(e(neqs)*numpar12-
e(tparms), lr8)

/*iv:estimate the VARs with 8 and 4 lags over the same sample period (1962Q2 - 2002Q1)*/
quietly var mt piet,lags(1/8) exog(d1 d2 d3)
scalar ldet88 = ln(e(detsig))
display ldet88
scalar numpar8 = e(tparms)/e(neqs)
quietly var mt piet in 10/169,lags(1/4) exog(d1 d2 d3)
scalar ldet4 = ln(e(detsig))
display ldet4

/*perform the likelihood ratio test for the null hypothesis of 4 lags*/
scalar lr4 = (e(T)-numpar8)*(ldet4-ldet88)
display "likelihood ratio = "= lr4
display "the significance level of the likelihood ratio stat = "= chi2tail(e(neqs)*numpar8-
e(tparms), lr4)

/*v:determine whether the seasonal dummy variables belong in the model*/
quietly var mt piet,lags(1/8)
scalar ldet80 = ln(e(detsig))
display ldet80
scalar lr80 = (e(T)-numpar8)*(ldet80-ldet88)
display "likelihood ratio = "= lr80
display "the significance level of the likelihood ratio stat = "= chi2tail(e(neqs)*numpar8-
e(tparms), lr80)
```

8. The results of the Granger causality tests are

F-Tests, Dependent Variable GM1

Variable	F-Statistic	Signif
GM1	13.5880	0.0000000
INF	2.7003	0.0085625

F-Tests, Dependent Variable INF

Variable	F-Statistic	Signif
GM1	1.5226	0.1544934
INF	13.8415	0.0000000

The results indicate that $gm1$ Granger causes inf. At conventional significance levels, inf does not Granger cause $gm1$.

The VAR treats seasonality using deterministic dummy variables. One issue might be the presence of a seasonal unit root in the $M1NSA$ series. The 2-variable VAR omits the presence of other potentially important influences on the variables. The other examples in the text suggest that real *GDP* and interest rates can have important influences on the money supply and inflation.

9. Most of Question 9 is answered in the Programming Manual. For part F, notice that *dlrgdp* shows no contemporaneous response to either *dlrm2* or *drs* shocks. Hence, *dlrgdp* must be causally prior to *dlrm2* and *drs*. Also notice that *dlrm*2 shows no contemporaneous response to *drs*. Hence, *dlrm*2 must be causally prior to *drs*. The scales in the figures are difficult to interpret because the variables have different units. Notice that real *GDP* and the money supply are expressed in logarithmic form while *drs* is the change in the short-term interest rate. As such, to compare the magnitudes of the various effects is it typical to standardize units so as to express the responses in terms of standard deviations.

Program for STATA Users

```
clear
cd "x:\New_data"  /*change this line if the dataset is saved elsewhere*/
insheet using money_dem.txt
gen obs = _n
tsset obs
gen lrgdp = ln(rgdp)
gen dlrgdp = d.lrgdp
gen price = gdp/rgdp
gen lrm2 = ln(m2/price)
gen dlrm2 = d.lrm2
gen drs = d.tb3mo

/*A:estimate a VAR with 12 lags of dlrgdp,dlrm2 and drs*/
set matsize 800
var dlrgdp dlrm2 drs, lags(1/12)
scalar ldet12 = ln(e(detsig))
scalar numpar12 = e(tparms)/e(neqs)
/*the results from the following formula for AIC and SBC are different from those reported in
Stata*/
display "AIC ="= e(T)*ldet12 + 2*e(tparms)
display "SBC ="= e(T)*ldet12 + e(tparms)*ln(e(T))

var dlrgdp dlrm2 drs in 14/169, lags(1/8)
scalar ldet8 = ln(e(detsig))
display "AIC ="= e(T)*ldet8 + 2*e(tparms)
display "SBC ="= e(T)*ldet8 + e(tparms)*ln(e(T))

/*c:perform the likelihood ratio test for the null hypothesis of 8 lags*/
scalar lr8 = (e(T)-numpar12)*(ldet8-ldet12)
display "likelihood ratio = "= lr8
display "the significance level of the likelihood ratio stat = "= chi2tail(e(neqs)*numpar12-
e(tparms), lr8)
```

```
/*D:show that drs is not block exogenous for dlrgdp and dlrm2*/
quietly var dlrgdp dlrm2 drs, lags(1/12)
scalar ldet3d = ln(e(detsig))
scalar numpar3d = e(tparms)/e(neqs)
quietly var dlrgdp dlrm2,lags(1/12)
scalar ldet2d = ln(e(detsig))
scalar lrd = (e(T)-numpar3d)*(ldet2d-ldet3d)
display "likelihood ratio = "= lrd
display "the significance level of the likelihood ratio stat = "= chi2tail(e(tparms), lrd)

/*E:Granger-causality test*/
quietly var dlrgdp dlrm2 drs, lags(1/12)
vargranger

/*F:impulse response functions*/
varirf set results1
varirf create var1, order(dlrgdp dlrm2 drs) step(13) replace set(results1)
varirf ctable (var1 dlrgdp dlrgdp oirf) (var1 dlrgdp dlrm2 oirf) (var1 dlrgdp drs oirf) (var1
dlrm2 dlrgdp oirf) (var1 dlrm2 dlrm2 oirf) (var1 dlrm2 drs oirf) (var1 drs dlrgdp oirf) (var1
drs dlrm2 oirf) (var1 drs drs oirf),step(11) noci
varirf cgraph (var1 dlrgdp dlrgdp oirf) (var1 dlrm2 dlrgdp oirf) (var1 drs dlrgdp oirf) (var1
dlrgdp dlrm2 oirf) (var1 dlrm2 dlrm2 oirf) (var1 drs dlrm2 oirf) (var1 dlrgdp drs oirf) (var1
dlrm2 drs oirf) (var1 drs drs oirf),lstep(0) ustep(11) cilines

/*G:estimate a near-VAR*/
constraint 1 [dlrgdp]:l.dlrgdp l2.dlrgdp l3.dlrgdp l4.dlrgdp l5.dlrgdp l6.dlrgdp l7.dlrgdp
l8.dlrgdp l9.dlrgdp l10.dlrgdp l11.dlrgdp l12.dlrgdp
constraint 2 [dlrm2]:l.dlrgdp l2.dlrgdp l3.dlrgdp l4.dlrgdp l5.dlrgdp l6.dlrgdp l7.dlrgdp
l8.dlrgdp l9.dlrgdp l10.dlrgdp l11.dlrgdp l12.dlrgdp
var dlrgdp dlrm2 drs, lags(1/12) constraint(1 2)
```

10. Question 9 suggests that is appropriate to estimate three-variable twelve-lag VAR for *dlrdgp*$_t$, *dlrm2*$_t$ and *drs*$_t$. Now suppose we want the contemporaneous relationships among the variables to be:

$$\begin{bmatrix} e_{mt} \\ e_{yt} \\ e_{rt} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ g_{21} & 1 & g_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{mt} \\ \varepsilon_{rt} \end{bmatrix}$$

where: e_y, e_{mt} and e_{rt} are the regression residual from the *dlrgdp*$_t$, *dlrm2*$_t$ and *drs*$_t$ equations, and ε_{yt}, ε_{mt} and ε_{rt} are the pure shocks (*i.e.*, the structural innovations) to *dlrgdp*$_t$, *dlrm2*$_t$ and

drs_t, respectively.

A. Provide a plausible economic interpretation of this set of restrictions.

Answer: The first restriction indicates that real *GDP* has a contemporaneous effect on the real money supply. The second indicates that real *GDP* is contemporaneously affected by all variables in the system. The third indicates that only the interest rate shocks have a contemporaneous effect on the interest rate. As it stands, this makes little economic sense. The programming manual actually uses

$$\begin{bmatrix} e_{yt} \\ e_{mt} \\ e_{rt} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ g_{21} & 1 & g_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{mt} \\ \varepsilon_{rt} \end{bmatrix}$$

In this way, the contemporaneous values of real *GDP* and the interest rate are unaffected by the other variables in the system. In contrast, the real money supply responds to all three variables. Notice that the system is overidentified.

Sample Program for STATA Users

```
clear
cd "x:\New_data"
insheet using money_dem.txt
gen obs = _n
tsset obs
gen lrgdp = ln(rgdp)
gen dlrgdp = d.lrgdp
gen price = gdp/rgdp
gen lrm2 = ln(m2/price)
gen dlrm2 = d.lrm2
gen drs = d.tb3mo

set matsize 800
matrix A = (1,0,0\.,1,.\0,0,1) /*define the matrix that shows the contemporaneous
relationships among the variables*/
svar dlrgdp dlrm2 drs, aeq(A) lags(1/12)
varirf set results1
varirf create var1,step(11) replace set(results1)
varirf ctable (var1 drs dlrgdp oirf) (var1 drs dlrm2 oirf) (var1 drs drs oirf),step(12) noci
varirf cgraph (var1 drs dlrgdp oirf) (var1 drs dlrm2 oirf) (var1 drs drs oirf),lstep(0)
ustep(11) cilines
```

CHAPTER 6

COINTEGRATION AND ERROR-CORRECTION MODELS

Lecture Suggestions

Figure 6.1 and Worksheet 6.1 illustrate the concept of cointegration. Worksheet 6.2 illustrates spurious regressions. You can use Figures M6-1 and M6-2 below for further emphasis. The first two panels in Figure M6-1 show 100 realizations of two independent unit root processes. The $\{y_t\}$ and $\{z_t\}$ sequences were constructed as:

$$y_t = 0.1 + y_{t-1} + \varepsilon_{yt}$$

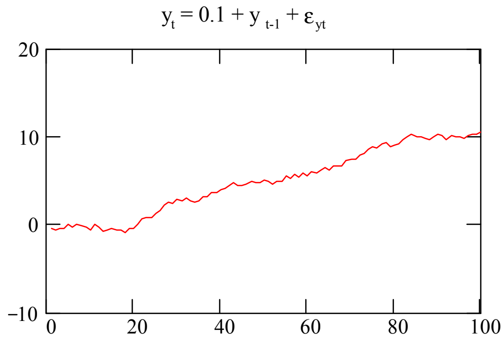
and:

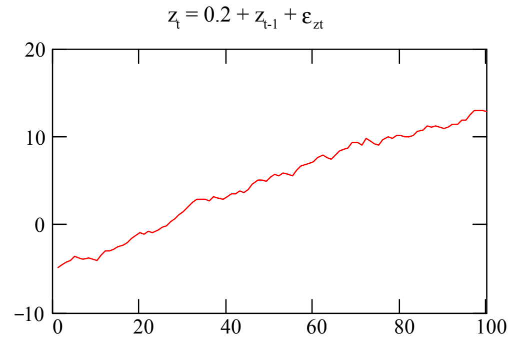
$$z_t = 0.2 + z_{t-1} + \varepsilon_{zt}$$

where: ε_{yt} and ε_{zt} are independent white-noise disturbances.

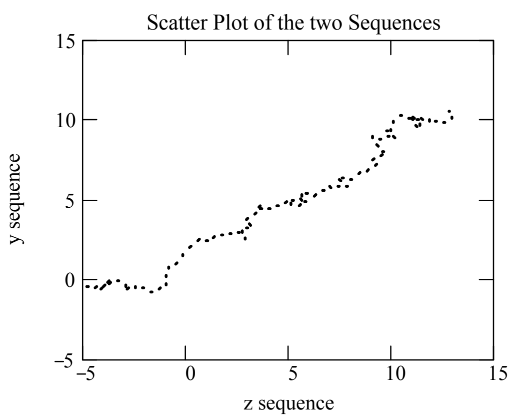
Two sets of one hundred random numbers were drawn to represent the $\{\varepsilon_{yt}\}$ and $\{\varepsilon_{zt}\}$ sequences. Using the initial values $y_0 = 0$ and $z_0 = -5$, the next 100 realizations of each were constructed using the formulas above. The drift terms impart a positive trend to each. Since each sequence tends to increase over time, the two appear to move together. The scatter plot in the third panel and the time plots in the fourth panel reflect this tendency. The spurious regression of y_t on z_t

Figure M6-1: A Spurious Regression

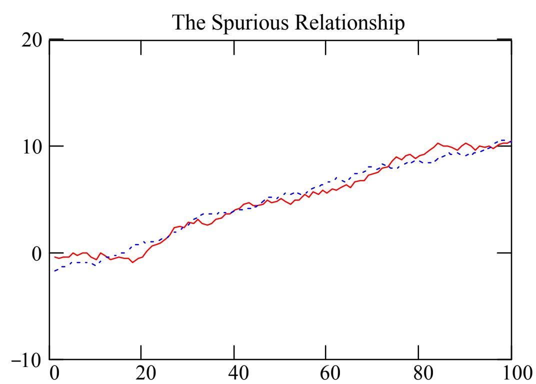




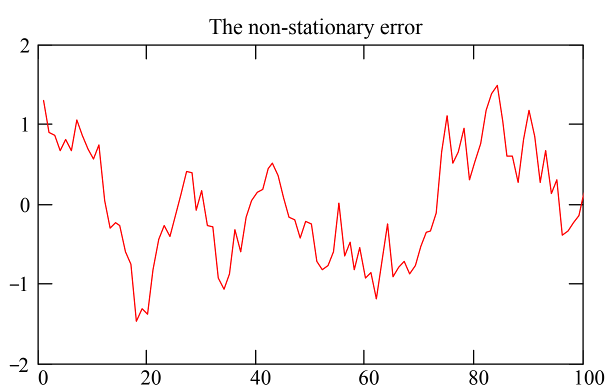
Both sequences were simulated as independent unit root processes.
Each has a positive drift so that the two sequences tend to increase over time. The relationship is spurious.



The scatter plotcaptures the tendency of both series to increase over time. The regression of y on z yields: $y = 0.68z + 1.63$.



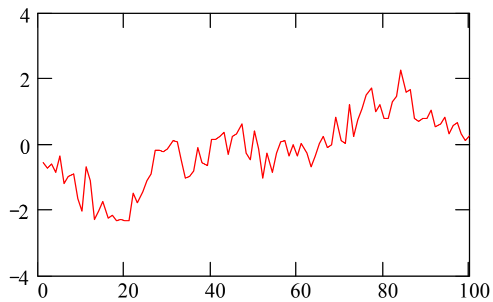
Transform the z sequence as: $w = 0.68z + 1.63$. The time paths of y and w seem to move together. However, the regression coefficients are meaningless.



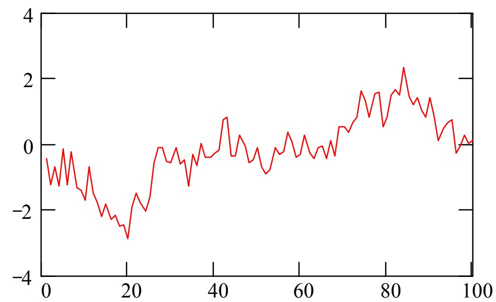
The regression error-term $y - 0.68z - 1.63$ is non-stationary. Hence, all deviations from the estimated relationship are permanent; the regression is spurious.

Figure M6-2: An Equilibrium Relationship

$y_t = y_{t\text{-}1} + \varepsilon_{yt}$

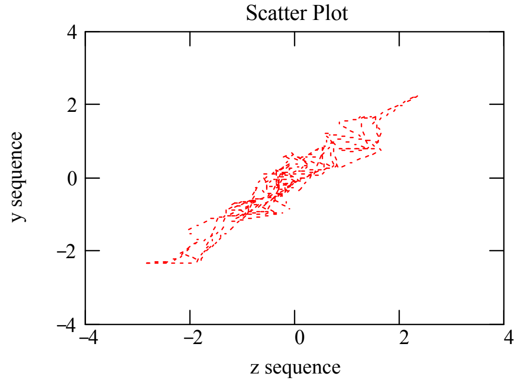


$z_t = z_{t\text{-}1} + \varepsilon_{zt}$

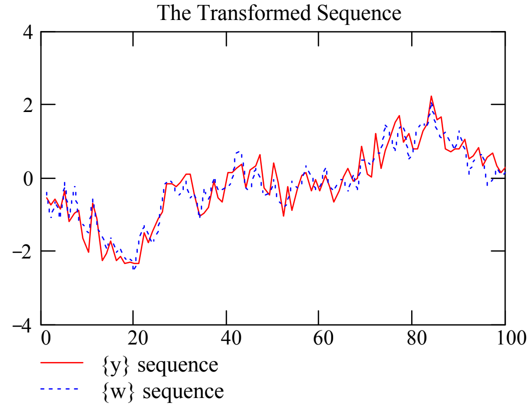


The simulated $\{y_t\}$ and $\{z_t\}$ sequences are both random-walk plus noise processes. Each meanders without any tendency to return to a long-run mean value. The error terms are: $\varepsilon_{yt} = \varepsilon_t + \eta_{yt} - \eta_{yt\text{-}1}$ and $\varepsilon_{zt} = \varepsilon_t + \eta_{zt} - \eta_{zt\text{-}1}$. Since each has the same stochastic trend, the $\{y_t\}$ and $\{z_t\}$ series are cointegrated.

Scatter Plot



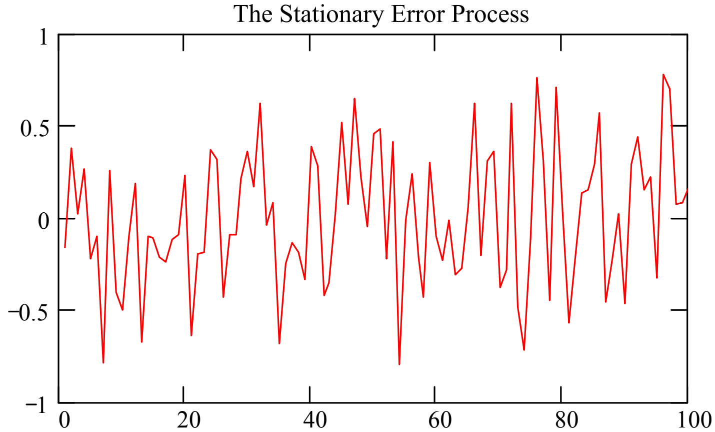
The Transformed Sequence



The scatter plot of z_t and y_t captures the tendency of both series to move together. The regression of y_t on z_t yields: $y_t = 0.889z_t + 0.007$.

Transform the z_t sequence as: $w_t = 0.889z_t + 0.007$. The time paths of y_t and w_t move together as a result of the common stochastic trend.

The Stationary Error Process



The regression error-term $y_t - 0.889z_t - 0.007$ is stationary. Hence, all deviations from the estimated relationship are temporary. The variables are cointegrated appears to have a "good" fit. However, regression coefficients are meaningless. The problem is that the error-term is a unit-root process; all deviations from the regression line are permanent.

In contrast, the simulated $\{y_t\}$ and $\{z_t\}$ sequences shown in Figure M6-2 are cointegrated. The two random-walk plus noise processes were simulated as:

$$y_t = y_{t-1} + \varepsilon_t + \eta_{yt} - \eta_{yt-1}$$
$$z_t = z_{t-1} + \varepsilon_t + \eta_{zt} - \eta_{zt-1}$$

where: ε_t, η_{yt}, and η_{zt} are computer generated random numbers.

The series have the same stochastic trend. The scatter plot in the third panel and the time plots in the fourth panel reflect the tendency of both to rise and fall together in response to the common $\{\varepsilon_t\}$ shocks. The regression of y_t on z_t yields a stationary error process. Hence, all deviations from the regression line are temporary.

Answers to Questions

1. Let equations (6.14) and (6.15) contain intercept terms such that:

$$y_t = a_{10} + a_{11}y_{t-1} + a_{12}z_{t-1} + \varepsilon_{yt} \qquad \text{and} \qquad z_t = a_{20} + a_{21}y_{t-1} + a_{22}z_{t-1} + \varepsilon_{zt}$$

A. Show that the solution for y_t can be written as:

$$y_t = [\ (1 - a_{22}L)\varepsilon_{yt} + (1-a_{22})a_{10} + a_{12}L\varepsilon_{zt} + a_{12}a_{20}]\ /\ [\ (1 - a_{11}L)(1 - a_{22}L) - a_{12}a_{21}L^2\ ]$$

Answer: Use lag operators to rewrite the system as:

$$(1 - a_{11}L)y_t - a_{12}Lz_t = a_{10} + \varepsilon_{yt}$$
$$-a_{21}Ly_t + (1-a_{22}L)z_t = a_{20} + \varepsilon_{zt}$$

Now write the system in matrix form as:

$$\begin{bmatrix} (1\ \ a_{11}L) & a_{12}L \\ a_{21}L & (1\ \ a_{22}L) \end{bmatrix}\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} a_{10} + \varepsilon_{yt} \\ a_{20} + \varepsilon_{zt} \end{bmatrix}$$

Using Cramer's Rule or matrix inversion, we can obtain the solutions for y_t as:

$$y_t = \frac{(1 a_{22}L)(a_{10} + \varepsilon_{yt}) + a_{12}L(a_{20} + \varepsilon_{zt})}{(1 a_{11}L)(1 a_{22}L)\ \ a_{12}a_{21}L^2}$$

Since $(1 - a_{22}L)a_{10} = (1-a_{22})a_{10}$, and $a_{12}La_{20} = a_{12}a_{20}$, the solution for y_t is verified.

B. Find the solution for z_t.

Answer: Use your answer to Part A and apply Cramer's Rule or matrix inversion to obtain:

$$\boxed{z_t = \frac{a_{21}L(a_{10}+\varepsilon_{yt})+(1\ \ a_{11}L)(a_{20}+\varepsilon_{zt})}{(1a_{11}L)(1a_{22}L)\ \ a_{12}a_{21}L^2}}$$

C. Suppose that y_t and z_t are $CI(1, 1)$. Use the conditions in (6.19), (6.20), and (6.21) to write the error-correcting model. Compare your answer to (6.22) and (6.23). Show that the error-correction model contains an intercept term.

Answer: Imposing the restrictions necessary to ensure that y_t and z_t are CI(1, 1), the equations for y_t and z_t can be written as:

$$\Delta y_t = -[a_{12}a_{21}/(1-a_{22})]y_{t-1} + a_{21}z_{t-1} + \varepsilon_{yt} + a_{10}$$
$$\Delta z_t = a_{21}y_{t-1} - (1 - a_{22})z_{t-1} + \varepsilon_{zt} + a_{20}$$

Normalizing the cointegrating vector with respect to y_{t-1}:

$$\Delta y_t = \alpha_y(y_{t-1} - \beta z_{t-1}) + \varepsilon_{yt} + a_{10}$$
$$\Delta z_t = \alpha_z(y_{t-1} - \beta z_{t-1}) + \varepsilon_{zt} + a_{20}$$

where: $\alpha_y = -a_{12}a_{21}/(1-a_{22})$; $\alpha_z = a_{21}$; and $\beta = (1-a_{22})/a_{21}$.

Thus, the error-correcting equations for Δy_t and Δz_t each contain a drift term. Another way to answer the question is to note that the solutions for y_t and z_t obtained in Parts A and B contain the deterministic expressions $[(1-a_{22})a_{10} + a_{12}a_{20}]$ and $[a_{21}a_{10} + (1-a_{11})a_{20}]$, respectively. Since the denominator contains a characteristic root equal to unity, the solution for each contains a deterministic trend.

D. Show that $\{y_t\}$ and $\{z_t\}$ have the same deterministic time trend (i.e., show that the slope coefficient of the time trends are identical).

Answer: The constant in the numerator of the solution for y_t is: $[(1-a_{22})a_{10} + a_{12}a_{20}]$. Since $1-a_{22} = a_{12}a_{21}/(1-a_{11})$, this constant can be rewritten as: $[a_{12}/(1-a_{11})][a_{21}a_{10} + (1-a_{11})a_{20}]$. Up to the expression $[a_{12}/(1-a_{11})]$, this deterministic numerator expression is the same as that in the solution for z_t. Given that the denominators are identical, y_t and z_t can be said to have the same deterministic time trend.

E. What is the condition such that the slope of the trend is zero? Show that this condition is such that the constant can be included in the cointegrating vector.

Answer: The $\{y_t\}$ sequence does not have a slope if $(1-a_{22})a_{10} + a_{12}a_{20} = 0$. Solving for a_{10} yields $a_{10} = -a_{12}a_{20}/(1-a_{22})$. Using this relationship, the error-correction equation for Δy_t is:

$$\Delta y_t = \alpha_y(y_{t-1} - \beta z_{t-1}) + \varepsilon_{yt} - a_{12}a_{20}/(1-a_{22})$$
$$= \alpha_y(y_{t-1} - \beta z_{t-1} + a_{20}/a_{21}) + \varepsilon_{yt}$$

Since $\alpha_z = a_{21}$, the error-correction model for Δz_t can be written as:

$\Delta z_t = \alpha_z(y_{t\text{-}1} - \beta z_{t\text{-}1} + a_{20}/a_{21}) + \varepsilon_{zt}$.

Thus, the normalized long-run equilibrium relationship is $y_{t\text{-}1} - \beta z_{t\text{-}1} + a_{20}/a_{21}$. The cointegrating vector has an intercept although the $\{\Delta y_t\}$ and $\{\Delta z_t\}$ sequences do not contain deterministic trends.

2. The data file **COINT6.PRN** contains the three simulated series used in sections 5 and 9. The following programs will reproduce the results.

Sample Program for RATS Users

```
all 100
open data a:coint6.prn              ;* The data disk is in drive a:\
data(format=prn,org=obs) / y z w
table                               ;* Produce summary statistics for y, z and w

set dy = y - y(t-1)                 ;* Take first-differences
set dz = z - z(t-1)
set dw = w - w(t-1)

linreg dy                           ;* Perform Dickey-Fuller test
# constant y{1}
linreg dy                           ;* Perform Augmented Dickey-Fuller test
# constant y{1} dy{1 to 4}

* Repeat the four lines above for z and w. Alternatively, you can use the procedure entitled
* DFUNIT.SRC. To use DFUNIT.SRC, type the statements
source c:\rats\dfunit.src       ;* The procedure is assumed the dfunit.src procedure is in the
                                :*RATS directory on drive c:
dfunit(lags=4) / y

linreg y / residy               ;* Estimate the long-run equilibrium relationship using y as
# constant z w                  ;*   the left-hand-side variable.  Save the residuals as "residy"

set dresidy = residy - residy{1}    ;* Obtain first-difference of the residuals
linreg dresidy                      ;* Perform the Dickey-Fuller test of the residuals
# residy{1}
linreg dresidy                      ;* Perform the Augmented Dickey-Fuller test
# residy{1} dresidy{1 to 4}
* Repeat the 7 lines above for z and w

system 1 to 3                   ;* Set up the system for the error-correction model
variables dy dz dw
```

```
lags 1 to 2                     ;* Use 2 lags of dy, dz, and dw
det constant residy{1}          ;* Include a constant and the error-correction term.  You can
end(system)                     ;*    use the residuals from the other two equilibrium relations
estimate(outsigma=vsigma)  * Estimate the model. Vsigma is the variance/covariance matrix

errors(impulses) 3 24 vsigma ;* Perform innovation accounting using the error-correction
# 1 ; # 2 ; # 3                 ;*    model


* To reproduce the results in Section 9, use the CATS procedure or the downloadable file
* entitled johansen.src. Note that the Johansen procedure in RATS does not allow you to use
* the variable name w. Redefine w using the following statement
set x = w
source c:\rats\johansen.src     ;*   It is assumed the johansen.src procedure is in the RATS
                                directory on drive c:\
@johansen.src(lags=2) /
# y z x
```

Note that johansen.src may inappropriately add seasonal dummy variables to your model. Moreover, there is no simple way to choose the form of the intercept term. If you use RATS, your answers will be slightly different from those reported in the text. For example, the λ_{max} and λ_{trace} statistics will be reported as:

```
lambda, lambda-max- and trace test
 0.32496   0.13401   0.02536
38.51272  14.10061   2.51767
 2.51767  16.61829  55.13101
```

Sample Program for STATA Users

```
clear
cd "x:\New_data"  /*change this line if the dataset is saved elsewhere*/
insheet using coint6.txt
summarize y z w
gen obs = _n
tsset obs

/*A: reproduce the results in section 5*/
/*step 1: pretest the variables for their order of integration*/
reg d.y l.y /*D-F regression of unit root test(no lags)*/
reg d.z l.z
reg d.w l.w
reg d.y l.y l.(d.y) l2.(d.y) l3.(d.y) l4.(d.y) /*augmented D-F regression of unit root test(four
lags)*/
reg d.z l.z l.(d.z) l2.(d.z) l3.(d.z) l4.(d.z)
```

```
reg d.w l.w l.(d.w) l2.(d.w) l3.(d.w) l4.(d.w)
/*Note:the command DFULLER is able to carry out all the estimations above as long as the
option 'regress' is included.*/

/*step 2: estimate the long-run equilibrium relationship*/
reg y z w
predict eyt,resid /*save the residuals to the series named eyt*/
reg z y w
predict ezt,resid
reg w y z
predict ewt,resid

/*check if eyt,ezt,ewt are stationary*/
reg d.eyt l.eyt
reg d.ezt l.ezt
reg d.ewt l.ewt
reg d.eyt l.eyt l.(d.eyt) l2.(d.eyt) l3.(d.eyt) l4.(d.eyt)
reg d.ezt l.ezt l.(d.ezt) l2.(d.ezt) l3.(d.ezt) l4.(d.ezt)
reg d.ewt l.ewt l.(d.ewt) l2.(d.ewt) l3.(d.ewt) l4.(d.ewt)

/*step 3: estimate the error-correction model*/
var d.y d.z d.w,lags(1) exog(l.ewt)

/*B: reproduce the results in section 9*/
/*Note: there are modules available to test for the number of cointegrating vectors in VARs
and estimate the vector error correction models after one or more cointegrating vectors have
been identified. They can be obtained by first searching "vector error correction model" in the
range "search all" and then installing the package called "VECECM"*/

johans y z w,lags(2) nonormal standard
johans y z w,lags(2) nonormal level(90)

/*the above two lines test the number of cointegrating vectors using Johansen procedure with
95% and 90% Osterwald-Lenum critical values respectively. In accord with the results in the
textbook, critical values of Case 1* with the assumption that the intercept term is in the
cointegration equation are appropriate*/
```

3. The file COINT_PPP.XLS contains quarterly values of German, Japanese, and Canadian wholesale prices and bilateral exchange rates with the United States. The file also contains the U.S. wholesale price level. The names on the individual series should be self-evident. For example, *p_us* is the U.S. price level and *ex_g* is the German exchange rate with the United States. All variables except the mark/dollar exchange rates run from 1973:Q4 to 2001:Q4 and all have been normalized to equal 100 in 1973:Q4.

A. Form the log of each variable. Estimate the long-run relationship between Canada and the United States as

$$\log(ex\_ca) = 4.12 + 0.937 \log(p\_ca) - 0.830\log(p\_us)$$

Do the point estimates of the slope coefficients seem to be consistent with long-run PPP?

Answer: Although the point estimates seem to be consistent with long-run PPP, you need to be a bit careful. There is a natural tendency to think that 0.937 is approximately equal to unity and 0.830 is approximately equal to minus one. However, inference on the cointegrating is unwarranted since the residuals from the regression are serially correlated and prices are not necessarily weakly exogeneous.

B. Since the residuals from the equilibrium regression contain a unit root, shocks to the real exchange rate never decay. Hence, long-run PPP fails.

C. **A RATS program that can perform the indicated tests is**

```
cal 1973 4 4                ;* The data set begins in 1973Q4 and ends in 2004Q4
all 2001:4
open data a:\coint_ppp.xls
data(org=obs,format=xls)

* Next, take the log of each variable
log ex_g / lex_g ; log ex_ca / lex_ca ; log ex_j / lex_j
log p_g / lp_g ; log p_j / lp_j ; log p_ca / lp_ca ; log p_us / lp_us

* You should now test each for a unit root
* The long-run relationship for the Canadian-U.S. rate can be obtained using
lin lex_ca / resids ; # constant lp_ca lp_us

* Now, test the residuals for a unit root
dif resids / dr
lin dr ; # resids{1} dr{1 to 3}

* Similarly, PPP for the German-U.S. rate can be tested using
lin lex_g / resids ; # constant lp_g lp_us
dif resids / dr
lin dr ; # resids{1} dr{1 to 4}
```

Sample STATA Program

```
clear
cd "x:\New_data" /*change this line if the dataset is saved elsewhere*/
```

```
insheet using coint_ppp.txt
gen obs = _n
tsset obs
gen lex_ca = ln(ex_ca)
gen lex_j = ln(ex_j)
gen ex_g1 = real(ex_g) /*note:some observations in the original series of ex_g are denoted by
the string "NA" so it is treated as a string variable. To perform estimation, however, ex_g
needs to be changed to a numeric variable by the "real" function.*/
gen lex_g1 = ln(ex_g1)
gen lp_ca = ln(p_ca)
gen lp_j = ln(p_j)
gen lp_g = ln(p_g)
gen lp_us = ln(p_us)

/*test if long-run PPP theory holds between Canada and US*/
reg d.lex_ca l.lex_ca l.(d.lex_ca) l2.(d.lex_ca) l3.(d.lex_ca) /*check if lex_ca is stationary*/
reg d.lp_ca l.lp_ca l.(d.lp_ca) l2.(d.lp_ca) l3.(d.lp_ca) /*check if lp_ca is stationary*/
reg d.lp_us l.lp_us l.(d.lp_us) l2.(d.lp_us) l3.(d.lp_us) /*check if lp_us is stationary*/
reg lex_ca lp_ca lp_us
predict mu1,residuals
reg d.mu1 l.mu1 l.(d.mu1) l2.(d.mu1) l3.(d.mu1)

/* The other countries can be tested in the same way*/
```

4. The second, fourth, and fifth columns of the file labeled INT_RATES.XLS contain the interest rates paid on U.S. 3-month, 3-year, and 10-year U.S. government securities. The data run from 1954:7 to 2002:12. These columns are labeled TBILL, r3, and r10, respectively.

RATS PROGRAM

```
cal 1954 7 12                 ;* The data set runs from July 1954 to December 2002
all 2002:12
open data a:\int_rates.xls
data(org=obs,format=xls)

*To test each series for a unit root using dfunit.src
source(noecho) c:\winrats\dfunit.src
@dfunit(ttest,lags=12) tbill
@dfunit(ttest,lags=12)  r3
@dfunit(ttest,lags=12)  r10

* The long-run relationship can be estimated using the T-bill rate as the 'dependent' variable.
linreg tbill / resids         ; * Save the residuals as resids
# constant r3 r10
```

```
* Perform the Engle-Granger test on resids
diff resids / dresids                  ;* Obtain first-difference of the residuals
linreg dresids                          ;* Perform the Dickey-Fuller test of the residuals
# resids1{1} dresids1{1 to 9}

* Repeat using the 10-year rate as the 'dependent' variable
linreg r10 / resids10
# constant r3 tbill
dif resids10 / dresids10
lin dresids10 ; # resids10{1} dresids10{1 to 4}     ; * Note the lag length of 4

* Note that some would use 12 lags
lin dresids10 ; # resids10{1} dresids10{1 to 12}

* To estimate the error-correction model you need to difference the variables
diff tbill / dtbill ; diff r3 / dr3 ; diff r10 / dr10

* Beginning with RATS 5.0, you can estimate the error-correction model as a system of
equations. Note that the resisuals from part B (i.e., resids) are used as the error-correction terms
system 1 to 3
variables dtbill dr3 dr10
lags 1 to 12
det resids{1}
end(system)
estimate(noftests,outsigma=v) / 1

* The multivariate AIC and SBC are calculated using
compute aic = %nobs * %logdet + 2*(38*3)
compute sbc = %nobs * %logdet + 38*3*log(%nobs)
display 'aic = ' aic 'sbc = ' sbc

* Now use only 6 lags
system 1 to 3
variables dtbill dr3 dr10
lags 1 to 6
det constant resids{1}
end(system)
estimate(noprint,noftests,outsigma=v) 1955:8 * 4
```

Sample Program for STATA Users

```
clear
cd "x:\New_data" /*change this line if the dataset is saved elsewhere*/
insheet using int_rates.txt
```

```
gen obs = _n
tsset obs

/*A:examine whether the variables have unit roots*/
dfuller tbill,lags(12)
display "estimated alpha1 ="=_b[l.tbill]
display "AIC ="=e(N)*ln(e(rss))+2*e(df_m)
dfuller tbill,lags(11)
display "AIC ="=e(N)*ln(e(rss))+2*e(df_m)
dfuller tbill,lags(10)
display "AIC ="=e(N)*ln(e(rss))+2*e(df_m)
dfuller tbill,lags(9)
display "AIC ="=e(N)*ln(e(rss))+2*e(df_m) /*AIC is increasing as the lag length decreases*/
dfuller r3,lags(12)
display "estimated alpha1 ="=_b[l.r3]
dfuller r10,lags(12)
display "estimated alpha1 ="=_b[l.r10]

/*B:estimate the cointegrating relationships using the Engle-Granger procedure*/
reg tbill r3 r10
predict resid1,residuals
dfuller resid1,lags(9)
display "estimated coefficient: "=_b[l.resid1]

/*C:use r10 as the dependent variable to redo Part B*/
reg r10 tbill r3
predict resid2,residuals
dfuller resid2,lags(4)
display "estimated coefficient: "=_b[l.resid2]
dfuller resid2,lags(12)
display "estimated coefficient: "=_b[l.resid2]

/*D:estimate an error-correction model using 12 lags of each variable*/
set matsize 800
var d.tbill d.r3 d.r10,lags(1/12) exog(l.resid1) noconstant
varlmar,mlag(12) /*check if the disturbances are not autocorrelated using LM test*/
/*obtain the three residual series and determine whether they appear to be white noise*/
predict residtb,equation(#1) residuals
wntestb residtb,table
wntestq residtb
predict residr3,equation(#2) residuals
wntestb residr3,table
wntestq residr3
```

```
predict residr10,equation(#3) residuals
wntestb residr10,table
wntestq residr10

/*E:estimate the model using Johansen procedure*/
/*Note: please install the package named "johans" first. It can be obtained by searching
"johans" in the range "all"*/
johans tbill r3 r10,lags(12) nonormal standard /*Case 1* is the appropriate table for critical
values in this model*/
```

5. Suppose you estimate π to be:

$$\pi = \begin{bmatrix} 0.6 & -0.5 & 0.2 \\ 0.3 & -0.25 & 0.1 \\ 1.2 & -1.0 & 0.4 \end{bmatrix}$$

A. Show that the determinant of π is zero.

Answer: Each element in row 1 is twice the corresponding element in row 2 (i.e., $\pi_{1i} = 2\pi_{2i}$) and half the corresponding element in row 3 (i.e., $\pi_{1i} = 0.5\pi_{3i}$). Since one row can be expressed as is a linear multiple of the another, the determinant must be zero.

B. Show that two of the characteristic roots are zero and that the third is 0.75.

Answer: Construct the determinant:

$$\begin{vmatrix} 0.6-\lambda & 0.5 & 0.2 \\ 0.3 & 0.25-\lambda & 0.1 \\ 1.2 & 1.0 & 0.4-\lambda \end{vmatrix} = 0.75\lambda^2 - \lambda^3$$

The three values of λ such that the determinant is zero are obtained by solving the equation: $\lambda^2(\lambda - 0.75) = 0$, so that:

$$\boxed{\lambda = 0, 0, \text{ and } 0.75}$$

C. Let $\beta' = (3\ \ {-2.5}\ \ 1)$ be the single cointegrating vector normalized with respect to x_3. Find the (3×1) vector α such that $\pi = \alpha\beta'$.

Answer: The problem is to select $\alpha = (\alpha_1\ \alpha_2\ \alpha_3)'$ such that $\pi = (\alpha_1\ \alpha_2\ \alpha_3)'(3\ \ {-2.5}\ \ 1)$. Hence,

$\alpha_1 = \pi_{11}/3 = \pi_{12}/(-2.5) = \pi_{13}/1$ so that:
$\alpha_1 = 0.6/3 = -0.5/(-2.5) = 0.2/1 = 0.2$.

Repeating for α_2 and α_3 yields:

$$\boxed{\alpha_1 = 0.2;\ \alpha_2 = 0.3/3 = 0.1;\ \alpha_3 = 1.2/3 = 0.4}$$

How would α change if you normalized β with respect to x_1?

Answer: Normalizing with respect to x_1 yields: $\beta' = (1, -5/6, 1/3)$. All values of α_i must be three times larger. Hence:

$$\boxed{\alpha_1 = 0.6;\ \alpha_2 = 0.3;\ \alpha_3 = 1.2}$$

D. Describe how you could test the restriction $\beta_1 + \beta_2 = 0$.

Answer: The appropriate method is discussed on page 395 of the text. Using the notation in the text, for $r = 1$ the appropriate test statistic is:

$$\boxed{T\,[\,\ln(1 - \lambda_1^*) - \ln(1 - \hat{\lambda}_1^*)\,]}$$

Now suppose you estimate π to be:

$$\pi = \begin{bmatrix} 0.8 & 0.4 & 0.0 \\ 0.2 & 0.1 & 0.0 \\ 0.75 & 0.25 & 0.5 \end{bmatrix}$$

E. Show that the three characteristic roots are: 0.0, 0.5, and 0.9.

Answer: Form the determinant

$$\begin{vmatrix} 0.8 - \lambda & 0.4 & 0.0 \\ 0.2 & 0.1 - \lambda & 0.0 \\ 0.75 & 0.25 & 0.5 - \lambda \end{vmatrix} = -0.45\lambda + 1.4\lambda^2 - \lambda^3$$

Solving for the values of λ such that the determinant is zero yields: $\lambda(\lambda^2 - 1.4\lambda + 0.45) = 0$ so that:

$$\boxed{\lambda = 0,\ 0.5,\ 0.9}$$

F. Select β such that:

$$\beta = \begin{bmatrix} 0.8 & 0.75 \\ 0.4 & 0.25 \\ 0.0 & 0.5 \end{bmatrix}$$

Find the (3×2) matrix α such that $\pi = \alpha\beta'$.

Answer: Here β is not normalized with respect to x_1, x_2 or x_3. Nevertheless, it is possible to find α. Consider:

$$\alpha = \begin{bmatrix} 1 & 0 \\ 0.25 & 0 \\ 0 & 1 \end{bmatrix}$$

It is straightforward to verify $\pi = \alpha\beta'$.

6. Suppose that x_{1t} and x_{2t} are integrated of orders 1 and 2, respectively. You are to sketch the proof

that any linear combination of x_{1t} and x_{2t} is integrated of order 2. Towards this end:
A. Allow x_{1t} and x_{2t} to be the random walk processes:

$$x_{1t} = x_{1t-1} + \varepsilon_{1t} \qquad \text{and} \qquad x_{2t} = x_{2t-1} + \varepsilon_{2t}$$

i) Given the initial conditions x_{10} and x_{20}, show that the solution for x_{1t} and x_{2t} have the form $x_{1t} = x_{10} + \Sigma\varepsilon_{1t-i}$ and $x_{2t} = x_{20} + \Sigma\varepsilon_{2t-i}$.
Answer: Let the index of summation run from $i = 0$ to t-1 so that: $x_{1t} = \varepsilon_{1t} + \varepsilon_{1t-1} + ... + \varepsilon_{11} + x_{10}$. This challenge solution must satisfy $x_{1t} = x_{1t-1} + \varepsilon_{1t}$. The issue is whether:

$$\varepsilon_{1t} + \varepsilon_{1t-1} + ... + \varepsilon_{11} + x_{10} = \varepsilon_{1t-1} + ... + \varepsilon_{11} + x_{10} + \varepsilon_{1t}.$$

Clearly, the two sides are identical for all possible realizations of the $\{\varepsilon_{1t}\}$ sequence. The same argument applies to x_{2t}. Hence the two solutions are:

$$\boxed{x_{1t} = \sum_{i=0}^{t-1} \varepsilon_{1t-i} + x_{10} \ \textit{and}\ x_{2t} = \sum_{i=0}^{t-1} \varepsilon_{2t-i} + x_{20}}$$

ii) Show that the linear combination $\beta_1 x_{1t} + \beta_2 x_{2t}$ will generally contain a stochastic trend.
Answer: The linear combination $\beta_1 x_{1t} + \beta_2 x_{2t}$ will contain the expression:

$$\boxed{\beta_1 \sum_{i=0}^{t-1} \varepsilon_{1t-i} + \beta_2 \sum_{i=0}^{t-1} \varepsilon_{2t-i}}$$

The sum of these two stochastic trends is also a stochastic trend. Simply define $\varepsilon_t = \beta_1\varepsilon_{1t} + \beta_2\varepsilon_{2t}$, so that the linear combination $\beta_1 x_{1t} + \beta_2 x_{2t}$ contains the stochastic trend $\Sigma\varepsilon_t$.

iii) What assumption is necessary to ensure that x_{1t} and x_{2t} are $CI(1, 1)$?
Answer: Suppose $\{\varepsilon_{1t}\}$ and $\{\varepsilon_{2t}\}$ are perfectly correlated such that $\varepsilon_{2t} = \beta\varepsilon_{1t}$ where β is a constant. Now, select values of β_1 and β_2 such that $\beta_1 + \beta_2\beta = 0$. The linear combination $\beta_1 x_{1t} + \beta_2 x_{2t}$ does not contain a stochastic trend since:

$$\beta_1 \sum_{i=0}^{t-1} \varepsilon_{1t-i} + \beta_2 \beta \sum_{i=0}^{t-1} \varepsilon_{1t-i} = 0$$

B. Now let x_{2t} be integrated of order 2. Specifically, let $\Delta x_{2t} = \Delta x_{2t-1} + \varepsilon_{2t}$. Given initial conditions for x_{20} and x_{21}, find the solution for x_{2t}. [You may allow ε_{1t} and ε_{2t} to be perfectly correlated].
Answer: One simple way to find x_{2t} is to define $z_t = \Delta x_{2t}$ so that $z_t = z_{t-1} + \varepsilon_{2t}$. Given the initial condition for z_1 (= x_{21} - x_{20}), the solution for z_t is:
$z_t = \varepsilon_{2t} + \varepsilon_{2t-1} + ... + \varepsilon_{22} + z_1$ so that:
$x_{2t} = x_{2t-1} + \varepsilon_{2t} + \varepsilon_{2t-1} + ... + \varepsilon_{22} + x_{21} - x_{20}$.

Now eliminate x_{2t-1} by iterating backwards:

$$x_{2t} = \varepsilon_{2t} + \varepsilon_{2t-1} + ... + \varepsilon_{22} + x_{21} - x_{20} + (\varepsilon_{2t-1} + \varepsilon_{2t-2} + ... + \varepsilon_{22} + x_{21} - x_{20} + x_{2t-2}).$$

$= \varepsilon_{2t} + 2\varepsilon_{2t\text{-}1} + ... + 2\varepsilon_{22} + 2(x_{21} - x_{20}) + x_{2t\text{-}2}$

Continue to iterate backwards to obtain:

$$x_{2t} = \varepsilon_{2t} + 2\varepsilon_{2t\text{-}1} + ... + 2\varepsilon_{22} + 2(x_{21} - x_{20}) + [\varepsilon_{2t\text{-}2} + \varepsilon_{2t\text{-}3} + ... + \varepsilon_{22} + (x_{21} - x_{20}) + x_{2t\text{-}3}]$$

$$= \varepsilon_{2t} + 2\varepsilon_{2t\text{-}1} + 3\varepsilon_{t\text{-}3} + ... + 3\varepsilon_{22} + 3(x_{21} - x_{20}) + x_{2t\text{-}4}$$

The solution is:

$$\boxed{x_{2t} = \varepsilon_{2t} + 2\varepsilon_{2t\text{-}1} + 3\varepsilon_{2t\text{-}2} + 4\varepsilon_{2t\text{-}3} + ... + (t - 1)\varepsilon_{22} + t\, x_{21} - (t - 1)\, x_{20}}$$

ii. Is there any linear combination of x_{1t} and x_{2t} that contains only a stochastic trend? Is there any linear combination of x_{1t} and x_{2t} that does not contain a stochastic trend?

Answer: No. Rewrite the solution for x_{2t} as:

$$x_{2t} = \sum_{i=0}^{t\text{-}2} \varepsilon_{2t\text{-}i} + \sum_{i=0}^{t\text{-}2} \varepsilon_{2t\text{-}i\text{-}1} + \sum_{i=0}^{t\text{-}2} \varepsilon_{2t\text{-}i\text{-}2} + ... + \varepsilon_{22} + t\, x_{21} - (t - 1)\, x_{20}$$

Even if ε_{1t} and ε_{2t} are perfectly correlated, the linear combination of the two must contain the following expression:

$$\sum_{i=0}^{t\text{-}2} \varepsilon_{2t\text{-}i\text{-}1} + \sum_{i=0}^{t\text{-}2} \varepsilon_{2t\text{-}i\text{-}2} + ... + \varepsilon_{22}$$

Thus, the linear combination contains a *double* summation of the $\{\varepsilon_{2t}\}$ sequence. No linear combination contains only a stochastic trend. The point is that all linear combinations are $I(2)$.

C. Provide an intuitive explanation for the statement: If x_{1t} and x_{2t} are integrated of orders d_1 and d_2 where $d_2 > d_1$, any linear combination of x_{1t} and x_{2t} is integrated of order d_2.

Answer: If $x_{2t} \sim I(d_2)$, its solution will contain a d_2-tuple summation of the $\{\varepsilon_{2t}\}$ sequence. However, the solution for x_{1t} will contain only a d_1-tuple summation. There is no linear combination of the two sequences that reduce the *order* of the summation in the solution for x_{2t}.

7. The Programming Manual that accompanies this text contain a discussion of nonlinear least squares and maximum likelihood estimation. If you have not already done so, download the manual and programs from the Wiley website.

A. Section 5.5 in Chapter 4 contains discussion of the problem of conducting inference on the parameters of a cointegrating vector. Execute Program 4.10. Why is it a problem that only 16.8% of the true values of β_1 lie within a 95% confidence interval?

B. How would you modify the program so as to generate the Engle-Granger critical values?

The question is answered in the Programming Manual using RATS.

Page 122: Cointegration

STATA Users can obtain the results for question 7 using:

```
clear
/*A:this is to rewrite program 4.10 in the program manual in stata*/
capture program drop coint
program define coint,rclass
   version 8.0
   drop _all
   args alpha1 alpha2 beta1
   set obs 150
   gen time = _n
   tsset time
   gen yt = invnorm(uniform())
   gen xt = invnorm(uniform())
   gen e1t = invnorm(uniform())
   gen e2t = invnorm(uniform())
   scalar success = 0
   forvalues i=1/149 {
          replace yt =(1-`alpha1')*yt[`i'] + (`alpha1'*`beta1')*xt[`i'] + e1t[`i'+1] if (_n>`i')
          replace xt =(1-`alpha2'*`beta1')*xt[`i'] + `alpha2'*yt[`i'] + e2t[`i'+1] if (_n>`i')
          }
   reg yt xt in 51/150
      if `beta1'>=(_b[xt]-1.96*_se[xt]) & `beta1'<=(_b[xt]+1.96*_se[xt])  {
         return scalar success = success + 1
         }
   return scalar betahat = _b[xt]
end

set seed 2002
simulate "coint 0.1 0.1 1" success = r(success) betahat = r(betahat),reps(2000) /*Note: If you
want to try another set of parameters, just change the three numbers in the double quotes,
which are alpha1,alpha2,beta1 respectively */
summarize betahat,detail
quietly summarize success
display "The percentage of success is:"=r(N)/20

/*B: The DF critical values of Engle-Granger(1987)*/
clear
capture program drop cicv
program define cicv,rclass
   version 8.0
   drop _all
   set obs 100
   gen time = _n
```

```
   tsset time
   gen e1t = invnorm(uniform())
   gen e2t = invnorm(uniform())
   gen yt = sum(e1t)
   gen xt = sum(e2t)
   reg yt xt
   predict ut,residuals
   reg d.ut l.ut, noconstant
   return scalar tphi = -(_b[l.ut]/_se[l.ut])  /*in the original paper, the model is d.ut=-
_b*(l.ut)+et*/
end

set seed 2002
simulate "cicv" tphi = r(tphi),reps(10000)
sort tphi
display tphi[9900] "  " tphi[9500] "  " tphi[9000]
```

8. Section 3 of Chapter 2 of the Programming Manual that comes with this text estimate the relationship between the long-term and short-term interest rate as

$$tb1yr_t = 0.698 + 0.0916tb3mo_t$$

A. Use the data set MONEYDEM.XLS to estimate the error-correction model. Use five lags of each variable.

B. As shown in the manual, you can use the error-correction model to obtain the impulse response functions with a Choleski decomposition.

C. In the equation for $\Delta tb1yr_t$, the coefficient on the error-correction term is –0.098 with a *t*-statistic equal to –0.427. Why is it possible to argue that $tb1yr_t$ is weakly exogenous. How can you model the long-term rate using the general-to-specific approach?

Sample Program for STATA Users

```
clear
cd "x:\New_data"  /*change this line if the dataset is saved elsewhere*/
insheet using money_dem.txt
gen obs = _n
tsset obs

/*A:estimate the error-correction model using five lags of each variable*/
gen tb1yr1 = real(tb1yr) /*Note:tb1yr has been treated as a string variable because the first
two observations have been denoted by "NA". To perform all kinds of estimation on it, it has
```

```
to be first transformed into a numeric variable. The function "real" does this and leaves the
first two as missing values*/
reg tb1yr1 tb3mo
predict resid,residuals
gen dtb1yr = d.tb1yr1
gen dtb3mo = d.tb3mo
var dtb1yr dtb3mo,lags(1/5) exog(l.resid)

/*B:obtain the impulse response functions*/
quietly var tb1yr1 tb3mo,lags(1/6)
capture varirf set results1,replace
varirf create ecm,order(tb1yr1 tb3mo) step(24)
varirf ctable (ecm tb1yr1 tb1yr1 oirf) (ecm tb1yr1 tb3mo oirf) (ecm tb3mo tb1yr1 oirf) (ecm
tb3mo tb3mo oirf),step(24) noci
varirf ctable (ecm tb1yr1 tb1yr1 fevd) (ecm tb3mo tb1yr1 fevd) (ecm tb1yr1 tb3mo fevd)
(ecm tb3mo tb3mo fevd),step(24) noci
varirf ograph (ecm tb1yr1 tb1yr1 oirf) (ecm tb1yr1 tb3mo oirf)
varirf ograph (ecm tb3mo tb1yr1 oirf) (ecm tb3mo tb3mo oirf)
```

CHAPTER 7

NONLINEAR TIME-SERIES MODELS

Lecture Suggestions

1. Once you abandon the linear framework, it is necessary to select a specific nonlinear alternative. Unfortunately, the literature does not provide a solid framework for this task. It is possible to estimate a series as a GAR, Bilinear, TAR, LSTAR, ESTAR, Markov switching, or ANN process. General tests for nonlinearity do not have a specific alternative hypothesis. Lagrange Multiplier tests generally accept a number of nonlinear alternatives. The issue can be illustrated by the estimate industrial production series beginning on page 419. Although a nonlinear model may be appropriate, the final TAR specification is a bit doubtful. My own view is that an underlying theoretical model should guide the model selection process. For example, in Section 7, a TAR model was used since theory suggests that low-terrorism states should be more persistent than high-terrorism states. To make the point, I rely heavily on Questions 1 and 2. Question 1 is designed to give the student practice in formulating a nonlinear model that is consistent with an underlying economic model. Question 2 asks the student to think about the nature of the nonlinearity that is suggested by any particular nonlinear estimation. In guiding the class discussion, you might want to make an overhead transparency of Figure M7.1 below.

2. The estimation of many nonlinear models requires the use of a software package with a programming language. Although the syntax explained in RATS Programming Manual may not be directly compatible with your software package, the logic will be nearly identical. As such, you can have your students read the following sections of the Programming Manual: Nonlinear Least Squares in Chapter 1.4, Do Loops in Chapter 3.1; If-Then-Else Blocks in Chapter 4.1, and Estimating a Threshold Autoregression (beginning on page 130).

Answers to Questions

1. Let p_A and p_M denote the price of cotton in Alabama and Mississippi, respectively. The price gap, or discrepancy, is p_A - p_M. For each part, present a nonlinear model that captures the dynamic adjustment mechanism given in the brief narrative.

   A. A large price gap (in absolute value) tends to be eliminated very quickly as compared to a small gap.

   Answer: Let the price gap in period t be defined as $g_t = p_{At} - p_{Mt}$. One characterization is ESTAR specification from page 401:

$$g_t = \alpha_0 + \alpha_1 g_{t-1} + (1 - \exp[-\gamma(g_{t-1} - c)^2])\,(\beta_0 + \beta_1 g_{t-1})$$

   B. The price gap is closed more quickly if it is positive than if it is negative.

   Answer: Again, there are many possible specifications. A TAR specification is

$$g_t = \begin{cases} \rho_1 g_{t-1} \ \ if \ \ g_{t-1} > 0 \\ \rho_2 g_{t-1} \ \ if \ \ g_{t-1} \le 0 \end{cases}$$

   If ρ_1 and ρ_2 are both positive such that $\rho_2 > \rho_1$, the positive gap is closed more quickly.

   C. It costs ten cents to transport between Alabama and Mississippi. Hence, a price discrepancy of less than ten cents will not be eliminated by arbitrage. However, fifty percent of any price gap exceeding ten cents will be eliminated within a period.

   Answer: A band-TAR specification is

$$\Delta g_t = \begin{cases} \rho_1 g_{t-1} \ \ if \ \ g_{t-1} > 0.10 \\ 0 \ if \ | g_{t-1} | < 0.10 \\ \rho_1 g_{t-1} \ \ if \ \ g_{t-1} < -0.10 \end{cases}$$

   Here, there is no change in the gap if g_{t-1} is less than \$0.10. Adjustment occurs only if the absolute value of the gap exceeds \$0.10.

   D. The value of p_A, but not the value of p_M, responds to a price gap.

   Answer: Here, there is no need to use a nonlinear mechanism since a linear specification can cature the dynamics. Consider

   $\Delta p_{At} = \alpha_1(p_{At-1} - p_{Mt-1})$

2. Draw the phase diagram for each of the following processes

   A. The GAR model: $y_t = 1.5y_{t-1} - 0.5\, y_{t-1}^3 + \varepsilon_t$

   Answer: Note that we can reparameterize the model as the NLAR(1) process: $y_t = [1.5 - 0.5(y_{t-1})^2]y_{t-1}$. As you should infer from Panel (a) of Figure M7.1, equilibrium positions (i.e., y_t

$= y_{t-1}$) occur at -1, 0 and +1. However, 0 is not a stable point since $[1.5 - 0.5(y_{t-1})^2]$ exceeds 1 in the neighborhood of $y_{t-1} = 0$. Moreover, the absolute value of the slope exceeds 1 when $| y_{t-1} | > 5^{0.5}$. Hence, the system will converge to 1 beginning with values $0 < y_{t-1} < 5^{0.5}$ and will converge to -1 beginning with values $-5^{0.5} < y_{t-1} < 0$.

B. The TAR model: $y_t = 1 + 0.5y_{t-1} + \varepsilon_t$ if $y_{t-1} > 2$ and $y_t = 0.5 + 0.75y_{t-1} + \varepsilon_t$ if $y_{t-1} \le 2$.

Answer: Panel (b) of Figure M7.1 indicates a break in the slope of the phase diagram at $y_{t-1} = 2$. There is a single stable 'equilibrium' position at $y_{t-1} = 2$. However, there is more persistence when $y_{t-1} < 2$.

C. The TAR model: $y_t = 1 + 0.5y_{t-1} + \varepsilon_t$ if $y_{t-1} > 0$ and $y_t = -1 + 0.5y_{t-1} + \varepsilon_t$ if $y_{t-1} \le 0$.

Answer: Panel (c) of Figure M7.1 indicates that this model is discontinuous at the threshold. Equilibrium positions occur when $y_{t-1} = +2$ and $y_{t-1} = -2$. Since the degree of auroregressive decay is always 0.5, both of the equilibrium values of the skeleton are stable.

D. The TAR model: $y_t = -1 + 0.5y_{t-1} + \varepsilon_t$ if $y_{t-1} > 0$ and $y_t = +1 + 0.5y_{t-1} + \varepsilon_t$ if $y_{t-1} \le 0$.

Answer: Panel (d) of Figure M7.1 indicates that this model is discontinuous at the threshold of zero. There is no value satisfying the relation $y_t = y_{t-1}$; hence, there is not an attractor for this model.

E. The LSTAR model: $y_t = 0.75y_{t-1} + 0.25y_{t-1}/[\ 1 + \exp(-y_{t-1})\ ] + \varepsilon_t$

Answer: Panel (e) of Figure M7.1 indicates that this model acts like a unit root process for large values of y_{t-1}. At the other extreme, only 75% of a highly negative value of y_{t-1} persists into the subsequent period.

F. The ESTAR model: $y_t = 0.75y_{t-1} + 0.25y_{t-1}[\ 1 - \exp(-\ y_{t-1}^2\ )\ ] + \varepsilon_t$

Answer: Panel (f) of Figure M7.1 indicates that this model acts like a unit root process when $| y_{t-1} |$ is large. When y_{t-1} is near zero, only 75% of y_{t-1} persists into the subsequent period.

3. In the Markov switching model, let p_1 denote the *unconditional* probability that the system is in regime one and let p_2 denote the *unconditional* probability that the system is in regime two. As in the text, let p_{ii} denote the probability that the system remains in regime i. Prove the assertion

$$p_1 = (1 - p_{22})/(2 - p_{11} - p_{22})$$
$$p_2 = (1 - p_{11})/(2 - p_{11} - p_{22})$$

Answer: Note that the probabilities are independent of time. Unconditional probability of being in regime 1 (p_1) is given by:

$p_1 = p_1p_{11} + p_2(1 - p_{22})$

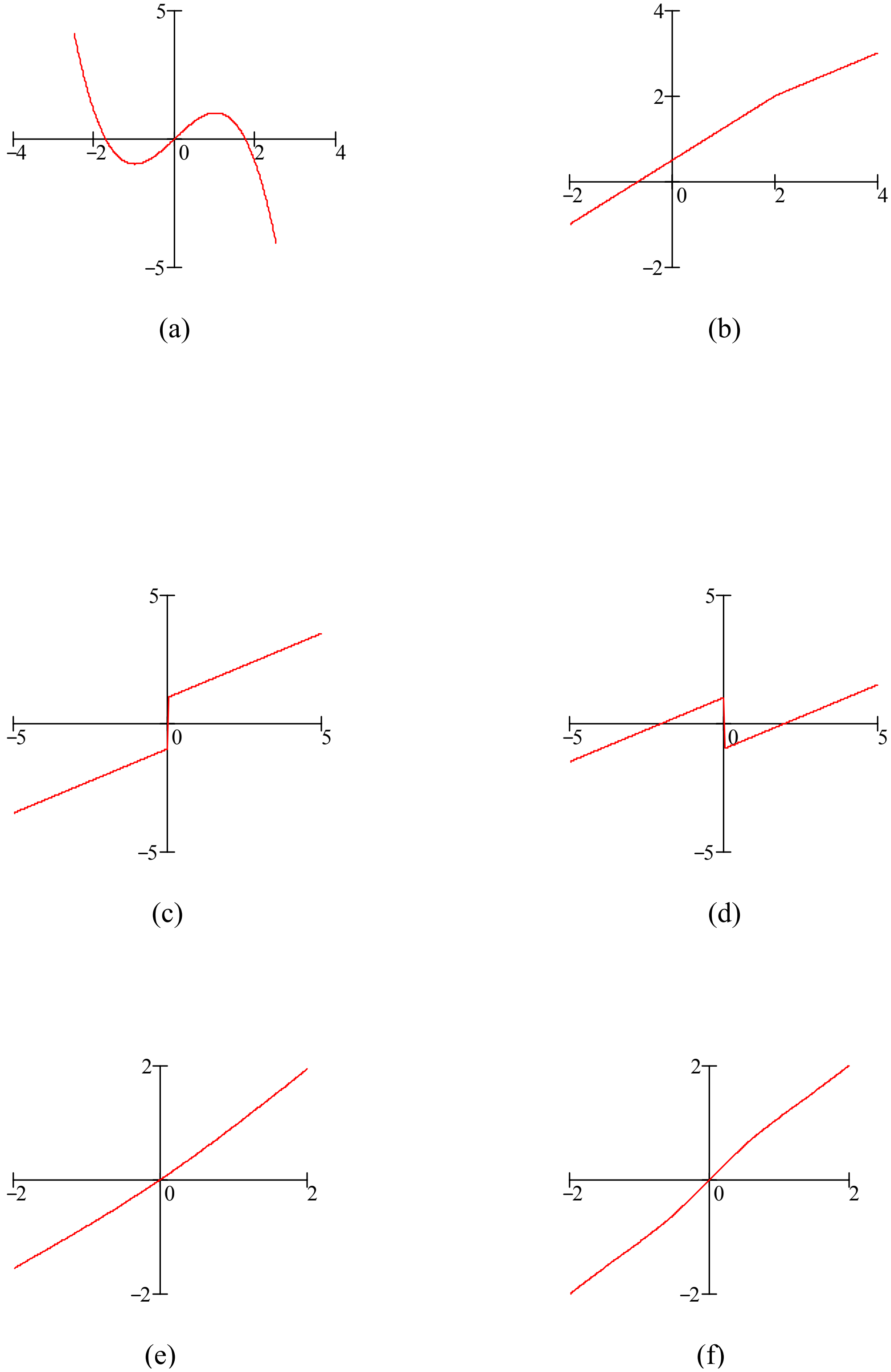
Note p_1 is expressed as a sum:

p_1p_{11} is the probability that the system is in regime 1 (p_1) multiplied by the probability of remaining in state 1 and remains in regime 1 (p_{11})

$p_2(1 - p_{22})$ is the probability that the system is in state 2 (p_2) multiplied by the probability of switching to state 1 from state 2 ($1 - p_{22}$)

Since $p_2 = 1 - p_1$, we get $p_1 = (1 - p_{22})/(2 - p_{11} - p_{22})$

Figure M7.1: The Six Phase Diagrams



4. The file labeled LSTAR.XLS contains the 250 realizations of the series used in Section 5.

Sample Program for RATS Users

```
all 250                                ;* The first three lines read in the data set
open data a:\lstar.xls
data(format=xls,org=obs)

* The correlations of {y_t} can be obtained using
cor(num=12,span=4,qstats) y

lin y / resids; # constant y{1}              ;* Estimate the linear model and save the residuals
                                             ;*  as the series resids
cor(number=24,span=4,qstats) resids          ;* Obtain the ACF of the residuals
set r2 = resids**2                           ;* Create the squared residuals
cor(number=24,span=4,qstats) r2              ;* The residuals and squared residuals show no
                                             ;* serial correlation
```

* B. Perform the RESET rest. To perform the RESET test we need to regress the residuals on
* the regressors and powers of the fitted values. Obtain the fitted values using
prj fitted   ; * The series *fitted* corresponds to $\{\hat{y}_t\}$. Next, obtain the powers of $\{\hat{y}_t\}$.
set f2 = fitted**2; set f3 = fitted**3 ; set f4 = fitted**4

* Next, regress the residuals on a constant, y_{t-1} and powers of the fitted values
lin resids ; # constant y{1} f2 f3 f4
exc ; # f2 f3 f4   ;* Test the exclusion of $(\hat{y}_t)^2$, $(\hat{y}_t)^3$ and $(\hat{y}_t)^4$
;* RATS users can also use the RESET.SRC procedure

* Now do the test using only $(\hat{y}_t)^2$ and $(\hat{y}_t)^3$
lin resids ; # constant y{1} f2 f3
exc ; # f2 f3   ;* Again, you reject the null of linearity

*D. Perform the LM test for LSTAR versus ESTAR adjustment. Let y_{t-1} be the delay
* parameter
set delay = y{1}

* To determine if y_{t-2} should be the delay parameter, use: set delay = y{2}

* Now, multiply y_{t-1} by the powers of delay parameter.
set y1d = y{1}*delay
set y1d2 = y{1}*delay**2
set y1d3 = y{1}*delay**3

* Regress *resids* on a constant, y_{t-1}, $(y_{t-1})^2$, $(y_{t-1})^3$, and $(y_{t-1})^4$
lin resids ; # constant y{1} y1d y1d2 y1d3

exc ; # y1d y1d2 y1d3   ;* Perform the *F*-test to see if you can reject the null of linearity

```
exc ; # y1d3                     ;* The t-test to see if you can reject the LSTAR in favor of the
                                 ;* ESTAR model

* Use nonlinear least squares to estimate gamma given that c = 4.4
com c = 4.4
nonlin a0 a1 b0 b1 gamma          ;* Instruct RATS to estimate α0, α1, β0 β1 and γ given c
* Next, create the LSTAR formula
frml lstar y = a0 + a1*y{1}  + (b0 + b1*y{1})/(1+exp(-gamma*(y{1}-c)))
* Next are the initial guesses
com a0 = 1.2, a1 = .5, b0 = 1. , b1 = -0.8, gamma = 20.
nlls(frml=lstar,iterations=100) y  2 250

* Now iterate. As illustrated in Table 7.1, you can use the estimated value of γ= 3.83 to find
* the estimates of α0, α1, β0 β1 and c. Use these estimated values to find an updated value of γ.

*/ E. If you estimate the process as a GAR process, you should find
```

$$y_t = 2.03 + 0.389y_{t-1} + 0.201y_{t-2} - 0.147\, y_{t-1}^2 + \varepsilon_t$$
$$(8.97) \quad (6.97) \quad (3.48) \quad (-10.57)$$

To estimate the GAR model, form the variable y_{t-1}^2

```
*/

     set y2 = y{1}**2
lin y / resids ; # constant y{1 to 2} y2 ; * Estimate the model and save the residuals
```

5. The file INDPROD_60.XLS contains the industrial production series used to estimate the TAR model in Section 6. Use the data to replicate the results reported in the text.

Sample Program for RATS Users

```
* The first four lines read in the data set. The sample runs from Jan. 1960 to June 2002
cal 1960 1 12
all 2002:6
open data a:\indpro_60.xls
data(org=obs,format=xls)

set y = 100*((indprod/indprod{1})-1)       ; *Create the growth rate

* Estimate the growth rate as an AR(3) and save the residuals as resids
lin y / resids; # constant y{1 to 3 }

* Obtain the ACF of the residuals and the squared residuals
cor(number=24,span=4,qstats) resids
set r2 = resids**2
cor(number=24,span=4,qstats) r2
```

```
* To perform the RESET test, obtain the fitted values
prj fitted

* Regress the residuals of the AR(3) on a constant, y_t-1 y_t-2, y_t-3 and powers of the fitted values
set f2 = fitted**2; set f3 = fitted**3 ;set f4 = fitted**4
lin resids ; # constant y{1 to 3 }  f2 f3 f4
exc ; # f2 f3 f4                                ;* Obtain the F-statistic

* Let d = 1 and suppose that the threshold is zero. Let flag denote the indicator variable. Set
* flag =1 if y_t-1 > 0
  set flag = %if(y{1}<0,0,1)

* Multiply y_t-1, y_t-2 and y_t-3 by flag
  set plus1 = flag*y{1}
  set plus2 = flag*y{2}
  set plus3 = flag*y{3}

* Multiply y_t-1, y_t-2 and y_t-3 by (1 - flag)
  set minus = 1 - flag
  set minus1 = (1-flag)*y{1}
  set minus2 = (1-flag)*y{2}
  set minus3 = (1-flag)*y{3}

* Regress y_t on flag, plus1, plus2, plus3, minus, minus1, minus2, and minus3
linreg y
# flag plus1 plus2 plus3 minus minus1 minus2 minus3

*The Programming Manual (see page 130) shows how to use Chan's method to obtain a
* consistent estimate of the threshold. Another example is found in Question 6 below.
```

6. The file GRANGER.XLS contains the interest rate series used to estimate the TAR and M-TAR models in Section 8.

RATS programmers can estimate the TAR and M-TAR models using

```
cal 1958 1 4                 ;* These four lines read in the data set. The data begin in
all 8 1994:1                 ;* 1958Q1 and end in 1994Q1
open data a:\granger.xls
data(format=xls,org=obs)

set spread = r_10 - r_short          ;* Create spread as the difference between the 10-year rate
dif spread / ds                      ;* and the short rate. Take the first difference of spread

* Perform the Dickey-Fuller test on the spread. Save the residuals as resids
lin ds / resids; # constant spread{1} ds{1}

* Perform the RESET test
```

```
prj fitted                                    ;* Obtain the fitted values
set f2 = fitted**2; set f3 = fitted**3        ;* Powers of the fitted values
lin resids ; # constant spread{1} ds{1} f2 f3 ;* Regress resids on a constant, spread(t-1)
                                              ;* and powers of the fitted values
exc ; # f2 f3                                 ;* F-test for excluding f2 and f3

* For Chan's method, we want exclude the lowest 15% and highest 15% of the observations
* from being used as  potential thresholds. One simple programming trick is to order the
* observations from lowest to the highest using
set thresh_test = spread ; order thresh_test
```

/* Now, *thresh_test* contains the sorted values of y_t. *Thresh_test*(1) is the smallest value of y_t and *thresh_test*(145) contains the largest value. Hence, thresh_test(22) through thresh_test(124) contain the desired middle 70% of the sorted observations. The first instruction below creates a variable *rss_test* that will be used to hold the residual sum of squares for the best fitting model. This value is initially set to be higher than any possible value of the estimated residual sum of squares. The second instruction creates the series that will hold the calculated residual sum of squares from each regression estimated. The third initializes the variable indicating the threshold (*thresh*) to be the first usable value.
*/

```
compute rss_test = 1000000.0
set rss = 0.
compute thresh = thresh_test(22)
```

/* Next, we begin the loop. For each value of *i* running from 22 to 124, we take the associated value of *thresh_test* and use it as a potential threshold. Inside the DO loop, the program creates the series *flag*. Note that *flag* =0 if $spread_{t-1}$ - *thresh_test*(*i*) [i.e., the potential threshold] is negative. Otherwise, *flag* = 1. Also created are the series *minus*, *spread_plus* and *spread_minus*. Hence, for each potential threshold, the LINREG instruction estimates a TAR equation in the form of (7.27) and (7.29). The residual sum of squares is compared to *rss_tes*t. If the resulting residual sum of squares exceeds this value, two instructions in brackets are not executed, the value of *i* is incremented by 1 and the loop is repeated. However, if %rss is lower than *rss_test* (i.e., if the residual sum of squares from the current regression is lower than any from the prior regressions) the bracketed instructions are executed. The value of *rss_test* is replaced by the value of %rss and the value of *thresh* is equated to the *thresh_test (i.e*., current value of the test threshold). Once the loop is completed, *thresh* will hold the value of the threshold that yields the lowest residual sum of squares.
*/

```
do i = 22,124
  set flag = %if(spread{1}-thresh_test(i).le.0.,0,1)
  set minus = 1 - flag
  set spread_plus = flag*(spread{1}-thresh_test(i))
  set spread_minus = minus*(spread{1}-thresh_test(i))

  lin(noprint) ds
  # spread_plus spread_minus ds{1}
```

```
  com rss(i) = %rss
  if %rss < rss_test {
     compute rss_test = %rss
     compute thresh = thresh_test(i)
  }
end do i
```

* Once the loop is exited, *thresh* contain the threshold with smallest residual sum of squares.

```
  dis ' Threshold = ' thresh  ;* Display the threshold
```

* Now use *thresh* = -0.27, to estimate the model

```
  set flag = %if(spread{1}-thresh.le.0.,0,1)
  set minus = 1 - flag
  set spread_plus = flag*(spread{1}-thresh)
  set spread_minus = minus*(spread{1}-thresh)
  lin ds
  # spread_plus spread_minus ds{1}
```

exc ; # spread_plus spread_minus ;* The *F*-test for the restriction $\rho_1 = \rho_2 = 0$
restrict 1 ;* The test for $\rho_1 = \rho_2$

```
  # 1 2
  # 1 -1 0
```

* Calculate and display the AIC and SBC. Note that the number of regressors should include
* the fact that τ was estimated.

```
compute aic = %nobs*log(%rss) + 2*(%nreg+1)
compute sbc = %nobs*log(%rss) + (%nreg+1)*log(%nobs)
display  'aic = ' aic ' sbc = ' sbc
```

* For the M-TAR model, the loop is written using

do i = 22,124
set flag = %if(ds{1}.le.0.,0,1) ;* Set the indicator according to whether $\Delta spread_{t-1} < 0$

```
  set minus = 1 - flag
  set spread_plus = flag*(spread{1}-thresh_test(i))
  set spread_minus = minus*(spread{1}-thresh_test(i))

  lin(noprint) ds
  # spread_plus spread_minus ds{1}

  com rss(i) = %rss
  if %rss < rss_test {
     compute rss_test = %rss
     compute thresh = thresh_test(i)
  }
```

```
end do i

* Once you exit the loop, estimate the M-TAR model using
  set flag = %if(ds{1}.le.0.,0,1)
  set minus = 1 - flag
  set spread_plus = flag*(spread{1}-thresh)
  set spread_minus = minus*(spread{1}-thresh)

  lin ds / resids
  # spread_plus spread_minus ds{1}

* The nonlinear error-correction model is estimated using spread_plus and spread_minus as
* the error correction terms.

dif r_10 / drl                    ;* Take the first difference of the two interest rates
dif r_short / drs

system 1 to 2
var drl drs
lags 1 to 2
det  spread_plus spread_minus
end(system)
estimate(outsigma=v)
```

7. Consider the linear process $y_t = 0.75y_{t-1} + \varepsilon_t$. Given $y_t = 1$, find $E_t y_{t+1}$, $E_t y_{t+2}$ and $E_t y_{t+3}$.

Now consider the GAR process $y_t = 0.75y_{t-1} - 0.25\, y_{t-1}^2 + \varepsilon_t$. Given $y_t = 1$, find $E_t y_{t+1}$. Can you find $E_t y_{t+2}$ and $E_t y_{t+3}$? [Hint: $(E_t y_{t+1})^2 \neq E_t(\, y_{t+1}^2\,)$].

Answer: In the linear case, we can use linear projections such that $E_t y_{t+1} = 0.75y_t$ and that $E_t y_{t+2} = E_t(E_{t+1}y_{t+2}) = E_t(0.75y_{t+1}) = 0.75E_t y_{t+1} = 0.75^2 y_t$. Since the expectations operator is linear, this method is not feasible in nonlinear equations. When $y_t = 1$, it follows that $E_t y_{t+1} = 0.5$. However, $E_t y_{t+2} = E_t[\, 0.75y_{t+1} - 0.25(y_{t+1})^2] = 0.75E_t(y_{t+1}) - 0.25\, E_t(\, y_{t+1}^2\,)$ and $E_t y_{t+3} = 0.75E_t y_{t+2} - 0.25E_t(\, y_{t+2}^2\,)$. As discussed in Section 7, it is necessary to use numerical integration to find the various values of $E_t(\, y_{t+k}^2\,)$.

SEMESTER PROJECT

The best way to learn econometrics is to estimate a model using actual data. At the beginning of the semester (quarter), students should identify a simple economic model that implies a long-run equilibrium relationship between a set of economic variables. Data collection should begin as early as possible so that the econometric tests can be performed as they are covered in class. Some students may be working on projects for which they have data. Others should be able to construct a satisfactory data set using the internet. Some of the web sites that were used in writing the test are

1. www.fedstats.gov/ The gateway to statistics from over 100 U.S. Federal agencies such as the Bureau of Labor Statistics, the Bureau of Economic Analysis, and the Bureau of the Census.

2. www.research.stlouisfed.org/fred2/ The St. Louis Fed Database. With over 1000 downloadable economic variables, this is probably the best site for economic time-series data.

3. www.nyse.com/marketinfo/marketinfo The New York Stock Exchange: The Data Library contains daily volumes and closing prices for the major indices

4. www.oecd.org/statistics/ The statistics portal for the Organization for Economic Cooperation and Development. Economic indicators, leading indicators, and labor force statistics.

For those who spend too much time searching for a project, students can update MONEY_DEM.XLS. This data set is used in the Programming Manual. The file contains quarterly values of seasonally adjusted U.S. nominal GDP, real GDP in 1996 dollars (RGDP), the money supply as measured by M2 and M3, and the 3-month and 1-year treasury bill rates for the period 1959:1 – 2001:1. Both interest rates are expressed as annual rates and the other variables are in billions of dollars. The data were obtained from the website of the Federal Reserve Bank of St. Louis and saved in Excel format.

The semester project is designed to employ all of the material covered in the text. Each student is required to submit a paper demonstrating competence in using the procedures. I require my students to use the format below. The various sections are collected throughout the semester so that student progress can be monitored. At the end of the semester, the individual sections are compiled into the final course paper. Of course, you might want to adapt the outline to your specific emphasis and to the statistical software package available to you. In Example 1, the student wants to estimate a demand for money function. In Example 2, the student wants to estimate the term structure of interest rates.

1. **Introduction**

Of course, it is important that students learn to generate their own research ideas. However, in the short space of a semester (or quarter), it is necessary for students to quickly select a semester project. After two weeks, I ask my students for a page or two containing:

1. A statement of the objective of the paper

2. A brief description of the relevant literature including the equation(s) to be estimated

3. The definition and source of each series to be used in the project. Some mention should be made concerning the relationship between the variables in the theoretical model and the actual data available.

Example 1: The student discusses why the demand for money can be represented by:

$m_t = \beta_0 + \beta_1 y_t + \beta_2 r_t + p_t + e_t$

where: m_t = money supply (=money demand), y_t = measure of income or output; r_t = vector of interest rates; p_t = price index; t is a time-subscript; e_t is an error-term; and all variables are measured in logarithms. Note that these variables are included on the data set MONEY_DEM.XLS.

Example 2: The student discusses why the term structure of interest rates implies a relationship among short-term and long-term interest rates of the form:

$TBILLt = \beta_0 + \beta_1 R3_t + \beta_2 R10_t + e_t$

where: $TBILL_t$ is the treasury bill rate, $R3_t$ is a three-year rate and $R10_t$ is a ten year rate. Note that these three variables are on the file INT_RATES.XLS.

2. **Difference equation models**

This portion of the project is designed to familiarize students with the application of difference equations to economic time-series data. Moreover, the initial data manipulation and creation of simple time-series plots introduces the student to the software at an early stage in the project. In this second portion of the paper, students should:

1. Plot the time path of each variable and describe its general characteristics. There should be some mention of the tendencies for the variables to move together.

2. For each series, develop a simple difference equation model that mimics its essential features.

Example 1: This portion of the paper contains time-series plots of the money supply, output, interest rate(s), and price index. The marked tendency for money, output and prices to steadily increase is noted. Periods of tranquility and volatility are indicated and the student mentions that the periods are similar for all of the variables. The student indicates that a difference equation with one or more characteristic root lying outside the unit circle might capture the time path of the money supply, price index, and level of output.

Example 2: This portion of the paper contains time-series plots of the various interest rate series. The tendency for the rates to meander is noted. The student shows how difference equations with a characteristic root that is unity can mimic the essential features of the interest rate series. It is also shown that characteristic roots near unity will impart similar time-paths to the series. The tendency for the rates to move together and any periods of tranquility and volatility are mentioned.

3. **Univariate properties of the variables**

This portion of the project introduces the student to the tools used in estimating the univariate properties of stationary time-series. Chapter 2 and Sections 1 through 7 of Chapter 3 provide the necessary background material. The student should select two or three of the key variables and for each:

1. Estimate an ARIMA model using the Box-Jenkins technique.

2. Provide out-of-sample forecasts. There should be some mention of the forecasting performance of the models.

3. Determine whether there are any GARCH and/or ARCH-M effects.

Example 1: The Box-Jenkins method suggests several plausible models for the money supply. Each is examined and compared in detail since the project focuses on money demand. In addition, there are periods of volatility that suggest GARCH errors. The GARCH estimates are presented.

Example 2: The three interest rate series are estimated using the Box-Jenkins methodology. The focus is not on any single interest rate series. Instead, several reasonable models for each series are found. Tests for the presence of GARCH and ARCH-M effects are presented.

4. **Estimates of the trend**

Most students will select one or more variables that exhibit evidence of non-stationary behavior. In this portion of the project, students should:

1. Use the material in the second portion of Chapter 3 to discuss plausible models for the trends. The goal is to refine the difference equation model suggested in Step 2 of the project.

2. Decompose the variables into their temporary and permanent components.

3. Use the material in Chapter 4 to conduct formal tests for unit roots and/or deterministic time trends. There can be a comparison of the effects of "detrending" versus differencing a series showing evidence of a trend.

4. Potentially important seasonal effects and/or evidence of structural change should be noted. If warranted, seasonal unit root tests and/or Perron tests for unit roots in the presence of structural breaks should be conducted.

Example 1: The money supply is decomposed into its temporary and permanent components using a Beveridge-Nelson decomposition. Dickey-Fuller tests for unit roots in the money supply series are conducted. The various tests for the presence of drift and/or time trends are also conducted. Given that the Federal Reserve has changed its operating procedures, the money supply is tested for unit roots in the presence of a structural break.

Example 2: The long-term rate is decomposed into its temporary and permanent components using a Beveridge-Nelson decomposition. Dickey-Fuller tests for unit roots in all of the interest rate series are conducted. The presence of unit roots in the interest rate series is mixed. It may be that the interest rates are near unit root processes. The student compares the ARMA estimates of the long-term rate using levels, "detrended" values, and first-differences of the data.

5. **Vector Autoregression Methods**

This portion of the project introduces multiple time-series methods. Chapter 5 of the text provides the background material necessary for the student to estimate a VAR. The student should complete the following tasks:

1. Estimate the variables as a VAR. The relationship among the variables should be analyzed using innovation accounting (impulse response functions and variance decomposition) methods.

2. Compare the VAR forecasts to the univariate forecasts obtained in Step 3 of the project.

3. A structural VAR using the Sims-Bernanke or Blanchard and Quah techniques should be attempted.

Example 1: The money supply, income level, interest rate and price level are estimated as an autoregressive system. It is reported that the univariate forecasts from Step 3 are nearly the same as those from the VAR. Granger causality tests are performed in order to pare down the model. A Choleski decomposition with various orderings is used to decompose the forecast error variances of the variables. Impulse response functions are used to examine the effects of the various shocks on the demand for money. The student estimates a structural VAR such that contemporaneous real income and interest rate shocks are unaffected by the other variables in the system.

Example 2: The student estimates the three interest rates as a VAR. Since the issue of stationarity is unclear, the student estimates the VAR in levels and in first-differences. A Choleski decomposition with various orderings is used to decompose the forecast error variances of the variables. Impulse response functions are used to show the effects of shocks to 10-year and 3-year rates on short-term rates. The student uses a bivariate VAR to decompose the 10-year interest rate into its temporary and permanent components.

6. **Cointegration**

This portion of the project introduces the concept of cointegration. Chapter 6 of the text provides the appropriate background. The student should:

1. Conduct Engle-Granger and Johansen tests for cointegration

2. Estimate the error-correction model. The error-correction model should be used to analyze the variables using innovation accounting techniques.

Example 1: For some specifications, the Engle-Granger and Johansen tests for cointegration will reveal a long-run equilibrium relationship among the variables. For other specifications and other sample periods, there is no credible money demand function. If the variables are not cointegrated, the error-correction model is not estimated since the variables are not cointegrated. Instead, the student discusses some of the credible reasons underlying the rejection of the theory as presented.

Example 2: The Engle-Granger and Johansen tests for cointegration reveal a long-run equilibrium relationship among the interest rates. The error-correction model is estimated. Innovation accounting is conducted; the results are compared to those reported in Step 5.

7. **Nonlinearity**

The last portion of the project uses nonlinear time-series models presented in Chapter 7. The student should:

1. Discuss a possible reason why a nonlinear specification might be plausible.

2. Conduct a number of tests that are capable of detecting nonlinearity.

3. Compare the linear estimates to the estimates from a nonlinear model.

Example 1: One reason why Engle-Granger and Johansen tests may fail is that they implicitly assume a linear adjustment mechanism. The student tests the variables for linear versus nonlinear behavior.

Example 2: The text suggests that interest rate spreads are nonlinear. An inverted yield curve is far less persistent than a situation when short-term rates are below long-term rates. A nonlinear model of the spread in estimated and compared to the linear model.