are more likely to be generated from model 1. If the $\{z_i\}$ series is serially correlated, you should perform the test with a robust t-statistic, such as that in the study by Newey and West (1987).

10. A MODEL OF THE INTEREST RATE SPREAD

The term "textbook example" is supposed to connote a very clear-cut illustration. If you are looking for a textbook example of the Box–Jenkins methodology, go back to Section 7 or turn to Question 11 at the end of this chapter. In practice, we rarely find a data series that precisely conforms to a theoretical ACF or PACF. This section is intended to illustrate some of the ambiguities that can be encountered when using the Box–Jenkins technique. These ambiguities may lead two equally skilled econometricians to estimate and forecast the same series using very different ARMA processes. Many view the necessity of relying on the researcher's judgment and experience as a serious weakness of a procedure that is designed to be scientific. Yet, if you make reasonable choices, you will select models that come very close to mimicking the actual data-generating process.

It is useful to illustrate the Box–Jenkins modeling procedure by estimating a quarterly model of the spread between a long-term and a short-term interest rate. Specifically, the interest rate spread (s_t) can be formed as the difference between the interest rate on 5-year U.S. government bonds and the rate on 3-month treasury bills. The data used in this section are the series labeled R5 and TBILL in the file QUARTERLY.XLS. Exercise 12 at the end of this chapter will help you to reproduce the results reported below.

Panel (a) of Figure 2.5 shows the spread over the period from 1960Q1 to 2012Q4. Although there are a few instances in which the spread is negative, the difference between long- and short-term rates is generally positive (the sample mean is 1.21). Notice that the series shows a fair amount of persistence in that the durations when the spread is above or below the mean can be quite lengthy. Moreover, there do not appear to be any major structural breaks (such as a permanent jump in the mean or variance) in that the dynamic nature of the process seems to be constant over time. As such, it is quite reasonable to suppose that the $\{s_t\}$ sequence is covariance stationary. In contrast, as shown in Panel (b), the first difference of the spread seems to be very erratic. As you will verify in Exercise 12, the Δs_t series has little informational content that can be used to forecast its future values. As such, it seems reasonable to estimate a model of the $\{s_t\}$ sequence without any further transformations. Nevertheless, because there are several large positive and negative jumps in the value of s_t, some researchers might want to transform it so as to diminish its volatility. A reasonable number of such shocks might indicate a departure from the assumption that the errors are normally distributed. Although a logarithmic or a square root transformation is impossible because some realizations of s_t are negative, one could dampen the series using $y_t = \log(s_t + 3)$. The point is that you should always maintain a healthy skepticism of the accuracy of your model since the behavior of the data-generating process may not fully conform to the underlying assumptions of the methodology.

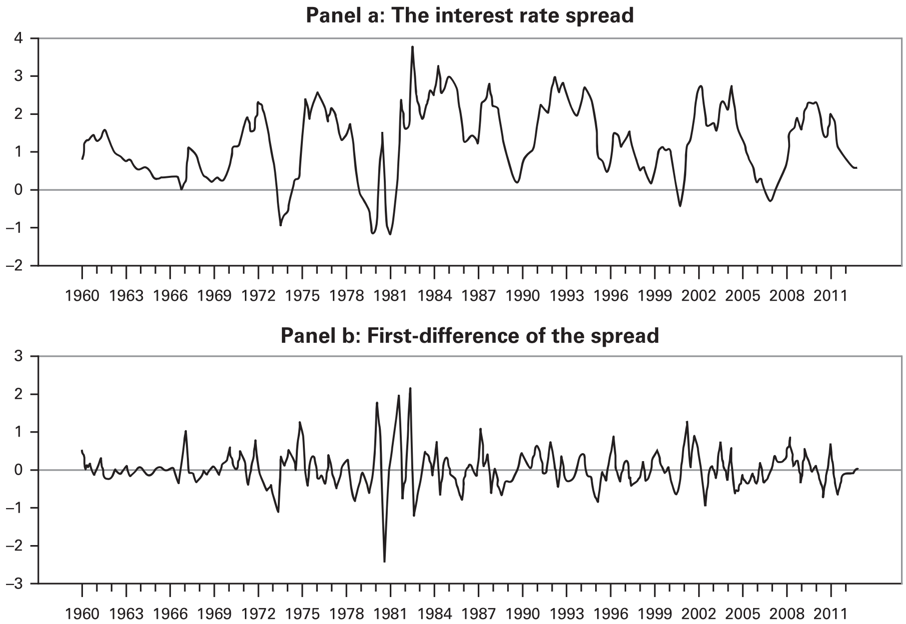

FIGURE 2.5 Time Path of the Interest Rate Spread

Before reading on, you should examine the autocorrelations and partial autocorrelation functions of the $\{s_t\}$ sequence shown in Figure 2.6. Try to identify the tentative models that you would want to estimate. Recall that the theoretical ACF of a pure MA(q) process cuts off to zero at lag q, and the theoretical ACF of an AR(1) model decays geometrically. Examination of Figure 2.6 suggests that neither of these

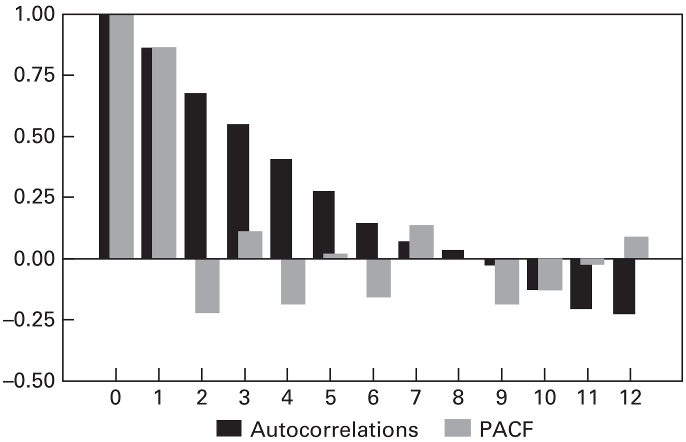

FIGURE 2.6 ACF and PACF of the Spread

specifications perfectly describes the sample data. In selecting your set of plausible models, also note the following:

1. The ACF and PACF converge to zero quickly enough that we do not have to worry about a time-varying mean. As suggested above, we do not want to *overdifference* the data and try to model the $\{\Delta s_t\}$ sequence.
2. The ACF does not cut to zero so that we can rule out a pure MA(q) process.
3. The ACF is not really suggestive of a pure AR(1) process in that the decay does not appear to be geometric. The value of ρ_1 is 0.857, and the values of ρ_2, ρ_3, and ρ_4 are 0.678, 0.550, and 0.411, respectively.
4. The estimated values of the PACF are such that $\phi_{11} = 0.858$, $\phi_{22} = -0.217$, $\phi_{33} = 0.112$, and $\phi_{44} = -0.188$. Although ϕ_{55} is close to zero, $\phi_{66} = -0.151$ and $\phi_{77} = 0.136$. Recall that, under the null hypothesis of a pure AR(p) model, the variance of $\phi_{p+i,p+i}$ is approximately equal to $1/T$. Since there are 212 total observations, the values of ϕ_{22}, ϕ_{44}, and ϕ_{66} are more than two standard deviations from zero (i.e., $2/212^{0.5} = 0.138$). In a pure AR(p) model, the PACF cuts to zero after lag p. Hence, if the s_t series follows a pure AR(p) process, the value of p could be as high as six or seven.
5. There appears to be an oscillating pattern in the PACF in that the first seven values alternate in sign. Oscillating decay of the PACF is characteristic of a positive MA coefficient.

Due to the number of small and marginally significant coefficients, the ACF and PACF of the spread are probably more ambiguous than most of those you will encounter. Hence, suppose you do not know where to start and estimate the s_t series using a pure AR(p) model. To illustrate the point, if you estimate the s_t series as an AR(7) process, you should obtain the estimates given in column 2 of Table 2.4. If you examine the table, you will find that all of the t-statistics on the first six lags exceed 1.96 in absolute value (indicating that the coefficients are significant the 5% level). Since t-statistic on the coefficient for y_{t-7} is 1.93, it is unclear as to whether to include the seventh lag. The sum or squared residuals (SSR) is 43.86 and the AIC and SBC are 791.10 and 817.68, respectively. The significance levels of the Q-statistics for lags 4, 8, and 12 indicate no remaining autocorrelation in the residuals.

Although the AR(7) model has some desirable attributes, one reasonable estimation strategy is to eliminate the seventh lag and estimate an AR(6) model over the same sample period. [Note that the data set begins in 1960Q1, so that with seven lags the estimation of the AR(7) begins in 1961Q4.] Although the autocorrelations of the residuals are such that $\rho_8 = 0.20$, the significance levels of the $Q(4)$, $Q(8)$, and $Q(12)$ statistics (equal to 0.29, 10.93, and 16.75) are 0.99, 0.21, and 0.16, respectively. As such, the Q-statistics suggest that you should not try to account to account for the residual autocorrelations at lag 8. Although a_5 appears to be statistically insignificant, it is generally not a good idea to use t-statistics to eliminate intermediate lags. As such, most researchers would not eliminate the fifth lag and estimate a model with lags 1 through 4 and lag 6. Recall that the appropriate use of a t-statistic requires that regressor in

Table 2.4 Estimates of the Interest Rate Spread

	AR (7)	AR (6)	AR (2)	$p = 1, 2, 7$	ARMA (1,1)	ARMA (2,1)	$p = 2$; $ma = (1, 7)$
a_0	1.20	1.20	1.19	1.19	1.19	1.19	1.20
	(6.57)	(7.55)	(6.02)	(6.80)	(6.16)	(5.56)	(5.74)
a_1	1.11	1.09	1.05	1.04	0.76	0.43	0.36
	(15.76)	(15.54)	(15.25)	(14.83)	(14.69)	(2.78)	(3.15)
a_2	−0.45	−0.43	−0.22	−0.20		0.31	0.38
	(−4.33)	(−4.11)	(−3.18)	(−2.80)		(2.19)	(3.52)
a_3	0.40	0.36					
	(3.68)	(3.39)					
a_4	−0.30	−0.25					
	(−2.70)	(−2.30)					
a_5	0.22	0.16					
	(2.02)	(1.53)					
a_6	−0.30	−0.15					
	(−2.86)	(−2.11)					
a_7	0.14			−0.03			
	(1.93)			(−0.77)			
β_1					0.38	0.69	0.77
					(5.23)	(5.65)	(9.62)
β_7							−0.14
							(−3.27)
SSR	43.86	44.68	48.02	47.87	46.93	45.76	43.72
AIC	791.10	792.92	799.67	801.06	794.96	791.81	784.46
SBC	817.68	816.18	809.63	814.35	804.93	805.10	801.07
$Q(4)$	0.18	0.29	8.99	8.56	6.63	1.18	0.76
$Q(8)$	5.69	10.93	21.74	22.39	18.48	12.27	2.60
$Q(12)$	13.67	16.75	29.37	29.16	24.38	19.14	11.13

Notes:
To ensure comparability, each equation was estimated over the 1961Q4 – 2012Q4 period.
Values in parentheses are the t-statistics for the null hypothesis that the estimated coefficient is equal to zero. SSR is the sum of squared residuals. $Q(n)$ are the Ljung–Box Q-statistics of the residual autocorrelations.
For ARMA models, many software packages do not actually report the intercept term a_0. Instead, they report the estimated mean of process, μ_y along with the t-statistic for the null hypothesis that $\mu_y = 0$. The historical reason for this convention is that it was easier to first demean the data and then estimate the ARMA coefficients than to estimate all values in one step. If your software package reports a constant term approximately equal to 0.216, it is reporting the estimated intercept.

question be uncorrelated with the other regressors. Given the autoregressive nature of the series, y_{t-5} is certainly correlated with y_{t-4} and y_{t-6}. The overall result is that the diagnostic checks of the AR(6) model suggest that it is adequate. In comparing the AR(6) and AR(7) models, the AIC selects the AR(7) model, whereas the SBC selects the more parsimonious AR(6) model.

Suppose that you try a very parsimonious model and estimate an AR(2). As you can see from the fourth column of the table, the AIC selects the AR(7) model, but SBC selects the AR(2) model. However, the residual autocorrelations from the AR(2) are problematic in that

ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6	ρ_7	ρ_8
0.03	−0.13	0.16	0.01	0.08	−0.10	−0.14	0.16

The Q-statistics from the AR(2) model indicate significant autocorrelation in the residuals at the shorter lags. As such, it should be eliminated from further consideration.

If you examined the AR(7) carefully, you might have noticed that a_3 almost offsets a_4 and that a_5 almost offsets a_6 (since $a_3 + a_4 \approx 0$ and $a_5 + a_6 \approx 0$). If you reestimate the model without s_{t-3}, s_{t-4}, s_{t-5}, and s_{t-6}, you should obtain the results given in column 5 of Table 2.4. Since the coefficient for s_{t-7} is now statistically insignificant, it might seem preferable to use the AR(2) instead. Yet, the AR(2) has been shown to be inadequate relative to the AR(7) and the AR(6) models.

Even though the AR(6) and AR(7) models perform relatively well, they are not necessarily the best forecasting models. There are several possible alternatives since the patterns of the ACF and PACF are not immediately clear. Results for a number of models with MA terms are shown in columns 6, 7, and 8 of Table 2.4:

1. From the decaying ACF, someone might try to estimate the ARMA(1, 1) model reported in column 6 of the table. The estimated value of a_1 (0.76) is statistically different from zero and is almost five standard deviations from unity. The estimated value of β_1 (0.38) is statistically different from zero and implies that the process is invertible. Notice that the SBC from the ARMA(1, 1) is smaller than that of the AR(7) and the AR(6). Nevertheless, the ARMA(1, 1) specification is inadequate because of remaining serial correlation in the residuals. The Ljung–Box Q-statistic for four lags of the residuals (equal to 6.63) has a significance level of 15.7%. As such, we cannot reject the null that $Q(4) = 0$ any conventional significance level. However, the $Q(8)$ and $Q(12)$ statistics indicate that the residuals from this model exhibit substantial serial autocorrelation. As such, we must eliminate the ARMA(1, 1) model from consideration.
2. Since the ACF decays and the PACF seems to oscillate beginning with lag 2 ($\phi_{22} = -0.217$), it seems plausible to estimate an ARMA(2, 1) model. As shown in column 6 of the table, the model is an improvement over the ARMA(1, 1) specification. The estimated coefficients ($a_1 = 0.43$ and $a_2 = 0.31$) are each significantly different from zero at conventional levels and imply characteristic roots in the unit circle. The AIC selects the ARMA(2, 1) model over that AR(6) and the SBC selects the ARMA(2, 1) over the AR(6) and the AR(7). The values for $Q(4)$, $Q(8)$, and $Q(12)$ indicate that the autocorrelations of the residuals are not statistically significant at the 5% level. Consider the ACF of the residuals:

ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6	ρ_7	ρ_8
0.01	0.01	−0.07	−0.02	−0.03	−0.08	−0.15	0.15

3. In order to account for the serial correlation at lag 7, it might seem plausible to add an MA term to the model at lag 7. As given in the last column of the table, all of the estimated coefficients are of high quality. In particular, the coefficient for β_7 has a t-statistic of -3.27. The estimated values of a_1 and a_2 are similar to those of the ARMA(2, 1) model. Again, the Q-statistics indicate that the autocorrelations of the residuals are not significant at conventional level. Both the AIC and SBC select the ARMA[2,(1,7)] specification over any of the other models. You can easily verify that the MA coefficient at lag 7 provides a better fit than an AR coefficient at lag 7 and that an ARMA[2,(1,8)] model is inadequate.

Although the ARMA[2,(1,7)] model appears to be quite reasonable, other researchers might have selected a decidedly different model. Consider some of the alternatives listed below.

1. **Parsimony versus Overfitting:** In Section 7, we examined the issue of fitting an MA coefficient at lag 16 to a true AR(2) process. If you reexamine the example, you can understand why some researchers shy away from estimating a model with long lags lengths that are disjoint from those of other periods. In the example of the spread, the problem with the ARMA(2, 1) model is that there was a small amount of residual autocorrelation around lag 7 or 8. The addition of the MA coefficient at lag 7 yielded a model with a better fit and remedied the serial correlation problem. However, is it really plausible that ε_{t-7} has a direct effect on the current value of the interest rate spread while lags 3, 4, 5, and 6 have no direct effects? In other words, do the markets for securities work in such a way that what happens 7 quarters in the past has a larger effect on today's interest rates than events occurring in the more recent past? Moreover, as you can verify by estimating the ARMA[2,(1,7)] model, the t-statistic for β_7 over the 1982Q1–2012Q4 period is equal to 0.60 and is not statistically significant. Notice that Panel (b) of Figure 2.5 suggests that the volatility of the spread in the late 1970s and early 1980s is not typical of the entire sample. It could be the case that the realizations from this period are anomalies that have large effects on the coefficient estimates and their standard errors. Thus, even though the AIC and SBC select the ARMA[2,(1,7)] model over the ARMA(2, 1) model, some researchers would prefer the latter.

   More generally, *overfitting* refers to a situation in which an equation is fit to some of the idiosyncrasies of present in a particular sample that are not actually representative of the data-generating process. In applied work, no data set will perfectly correspond to every assumption required for the Box–Jenkins methodology. Since it is not always clear which characteristics of the sample are actually present in the data-generating process, the attempt to expand a model so as to capture every feature of the data may lead to overfitting.

2. **Volatility:** Given the volatility of the $\{s_t\}$ series during the late 1970s and early 1980s, transforming the spread using some sort of a square root or logarithmic transformation might be appropriate. Moreover, the s_t series has a

number of sharp jumps, indicating that the assumption of normality might be violated. For a constant c such that $s_t + c$ is always positive, transformations such as $\ln(s_t + c)$ or $(s_t + c)^{0.5}$ yield series with less volatility than the s_t series itself. Alternatively, it is possible to model the difference between the log of the 5-year rate and the log of the 3-month rate.

A general class of transformations was proposed by Box and Cox (1964). Suppose that all values of $\{y_t\}$ are positive so that it is possible to construct the transformed $\{y_t^*\}$ sequence as

$$\begin{aligned} y_t^* &= (y_t^\lambda - 1)/\lambda \qquad && \lambda \neq 0 \\ &= \ln(y_t) && \lambda = 0 \end{aligned}$$

The common practice is to transform the data using a preselected value of λ. The selection of a value for λ that is close to zero acts to "smooth" the sequence. An ARMA model can be fitted to the transformed data. Although some software programs have the capacity to simultaneously estimate λ along with the other parameters of the ARMA model, this approach has fallen out of fashion. Instead, it is possible to actually model the variance using the methods discussed in Chapter 3.

3. **Trends:** Suppose that the span of the data had been somewhat different in that the first observation was for 1973Q1 and the last was for 2004Q4. If you examine Panel (a) of Figure 2.4, you can see that someone might be confused and believe that the data contained an upward trend. Their misinterpretation of the data might be reinforced by the fact that the ACF converges to zero rather slowly. As such, they might have estimated a model of the Δs_t series. Others might have detrended the data using a deterministic time trend.

Out-of-Sample Forecasts

We can assess the forecasting performance of the AR(7) and ARMA[2,(1,7)] models by examining their bias and mean square prediction errors. Given that the data set contains a total of 205 (i.e., 205 = 212 − 7) usable observations, it is possible to use a holdback period of 50 observations. This way, there are at least 155 observations in each of the estimated models and an adequate number of out-of-sample forecasts. First, the two models were estimated using all available observations through 2000Q2 and the two one-step-ahead forecasts were obtained. The actual value of $s_{2000:3} = 0.40$; the AR(7) predicted a value of 0.697, and the ARMA[2,(1,7)] model predicted a value of 0.591. Thus, the forecast of the ARMA[2,(1,7)] is superior to that of the ARMA(7) for this first period. An additional 49 forecasts were obtained for periods 2000Q4 to 2012Q4. Let e_{1t} denote the forecast errors from the AR(7) model and e_{2t} denote the forecast errors from the ARMA[2,(1,7)] model. The mean of e_{1t} is 1.239, the mean of e_{2t} is 1.244, and the estimated variances are $\text{var}(e_1) = 0.797$ and $\text{var}(e_2) = 0.780$. As such, the bias of AR(7) is slightly smaller while the ARMA[2,(1,7)] has the smallest MSPE.

To ascertain whether these differences are statistically significant, we first check the bias. Let the $\{f_{1t}\}$ series contain the 50 forecasts of the AR(7) model and let $\{f_{2t}\}$

contain the 50 forecasts from the ARMA[2,(1,7)] model. Beginning with $t = 2000Q3$, we can estimate the two regression equations:

$$s_t = 0.0594 + 0.968f_{1t} \qquad \text{and} \qquad s_t = 0.004 + 1.004f_{2t}$$

For the AR(7) model, the F-statistic for the restriction that the intercept equals zero and the slope equals unity is 0.110 with significance level of 0.896. Clearly, the restriction of unbiased forecasts does not appear to be binding. For the ARMA[2,(1,7)] model, the F-statistic is 0.014 with a significance level of 0.986. Hence, there is strong evidence that both models have unbiased forecasts.

Next, consider the Granger–Newbold test for equal mean square prediction errors. Form the x_i and z_i series as $x_i = e_{1i} + e_{2i}$ and $z_i = e_{1i} - e_{2i}$, respectively. The correlation coefficient between x_i and z_i is $r_{xz} = 0.234$. Given that there are 50 observations in the holdback period, form the Granger–Newbold statistic

$$r_{xz}/\sqrt{(1 - r_{xz}^2)/(H - 1)} = 0.234/\sqrt{(1 - (0.234)^2)/49} = 1.69$$

With 49 degrees of freedom, a value of $t = 1.69$ is not statistically significant. We can conclude that the forecasting performance of the AR(7) is not statistically different from that of the ARMA[2,(1,7)].

Since the e_{1i} and e_{2i} series contain only a low amount of serial correlation, we obtain virtually the same answer using the DM statistic. Oftentimes, forecasters are concerned about the MSPE. However, there are many other possibilities. In Exercise 12 at the end of this chapter, you will be asked to use the mean absolute error. Now, to illustrate the use of the DM test, suppose that the cost of a forecast error rises extremely quickly in the size of the error. In such circumstances, the loss function might be best represented by the forecast error raised to the fourth power. Hence,

$$d_i = (e_{1i})^4 - (e_{2i})^4 \tag{2.63}$$

The mean value of the $\{d_i\}$ sequence from (2.63) (i.e., $\overline{d}$) is 0.01732, and the estimated variance is 0.002466. Since $H = 50$, we can form the DM statistic

$$\text{DM} = 0.01732/(0.002466/49)^{1/2} = 2.441$$

The null hypothesis is that the models have equal forecasting accuracy, and the alternative hypothesis is that the forecast errors from the AR[2,(1,7)] are smaller than those of the AR(7). With 49 degrees of freedom, the t-value of 2.441 is significant at the 1.829% level. Hence, there is evidence in favor of the AR[2,(1,7)] model. If there is serial correlation in the $\{d_t\}$ series, we need to use the specification in (2.63). Toward this end, we would want to select the statistically significant values of γ_q. The autocorrelations of d_t are

ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6	ρ_7	ρ_8	ρ_9	ρ_{10}	ρ_{11}	ρ_{12}
−0.10	−0.15	0.26	0.01	0.36	0.00	−0.09	0.13	0.06	0.05	−0.08	0.07

$$Q(4) = 5.53; Q(8) = 14.76; \text{and } Q(12) = 15.93$$

Although ρ_5 is large, many applied econometricians would dismiss it as spurious. It does not seem plausible that correlations for ρ_1 and ρ_2 are actually very close to zero while the correlation between d_t and d_{t-5} is very large. Moreover, the Ljung–Box $Q(4)$, $Q(8)$, and $Q(12)$ statistics do not indicate that the autocorrelations are significant. The significance levels are 0.237, 0.064, and 0.195, respectively. Nevertheless, if you do estimate the long-run variance using (2.63) with five lags, you should find that DM = 1.848 (so that the MSPEs are not statistically different from each other). The example underscores the point made earlier that there is no clear answer as to the best way to measure the long-run variance of $\overline{d}$ in the presence of serial correlation. The more general result is that the two models are not substantially different from each other. Both should provide reasonable forecasts.

11. SEASONALITY

Many economic processes exhibit some form of seasonality. The agricultural, construction, and travel sectors have obvious seasonal patterns resulting from their dependence on the weather. Similarly, the Thanksgiving-to-Christmas holiday season has a pronounced influence on the retail trade. In fact, the seasonal variation of a series may account for the preponderance of its total variance. Forecasts that ignore important seasonal patterns will have a high variance.

Too many people fall into the trap of ignoring seasonality if they are working with **deseasonalized** or **seasonally adjusted** data. Suppose you collect a data set that the U.S. Census Bureau has "seasonally adjusted" using its X–11, X–12, or X–13 methods.[5] In principle, the seasonally adjusted data should have the seasonal pattern removed. However, caution is necessary. Although a standardized procedure may be necessary for a government agency reporting hundreds of series, the procedure might not be best for an individual wanting to model a single series. Even if you use seasonally adjusted data, a seasonal pattern might remain. This is particularly true if you do not use the entire span of data; the portion of the data used in your study can display more (or less) seasonality than the overall span. There is another important reason to be concerned about seasonality when using deseasonalized data. Implicit in any method of seasonal adjustment is a two-step procedure. First, the seasonality is removed, and second, the autoregressive and moving average coefficients are estimated using Box–Jenkins techniques. As surveyed in Bell and Hillmer (1984), often the seasonal and the ARMA coefficients are best identified and estimated jointly. In such circumstances, it is wise to avoid using seasonally adjusted data.

Models of Seasonal Data

The Box–Jenkins technique for modeling seasonal data is only a bit different from that of nonseasonal data. The twist introduced by seasonal data of period s is that the seasonal coefficients of the ACF and PACF appear at lags $s, 2s, 3s, \ldots$, rather than at lags 1, 2, 3, ... For example, two purely seasonal models for quarterly data might be

$$y_t = a_4 y_{t-4} + \varepsilon_t, \qquad |a_4| < 1 \tag{2.64}$$

MA(1). (Note: You will need to transform your forecasts to the forecasts of the $cpicore_t$.)

14. The file QUARTERLY.XLS contains U.S. interest rate data from 1960Q1 to 2012Q4. As indicated in Section 10, form the spread by subtracting the T-bill rate from the 5-year rate.

a. Use the full sample period to obtain estimates of the AR(7) and the ARMA(1, 1) model reported in Section 10.

b. Estimate the AR(7) and ARMA(1, 1) models over the period 1960Q1–2000Q3. Obtain the one-step-ahead forecast and the one-step-ahead forecast error from each. As in Section 10, continue to update the estimation period so as to obtain the 50 one-step-ahead forecast errors from each model. Let f_{1t} denote the forecasts from the AR(7) and f_{2t} denote the forecasts from the ARMA(1, 1). You should find that the properties of the forecasts are such that

$$y_{2000Q3+t} = 0.0536 + 0.968 f_{1t} \text{ and } y_{2000Q3+t} = -0.005 + 1.000 f_{2t}.$$

Are the forecasts unbiased?

c. Construct the Diebold–Mariano test using the mean absolute error. How do the results compare to those reported in Section 10.

d. Use the Granger–Newbold test to compare the AR(7) model to the ARMA(1, 1).

e. Construct the ACF and PACF of the first difference of the spread. What type of model is suggested?

f. Show that a model with 2 AR lags and MA lags at 3 and 8 has a better fit than any of the models reported in the text. What do you think about such a model?

15. The file QUARTERLY.XLS contains the U.S. money supply as measured by M1 (M1NSA) and as measured by M2 (M2NSA). The series are quarterly averages over the period 1960:1 to 2012Q4.

a. Reproduce the results for M1 that are reported in Section 11 of the text.

b. How do the three models of M1 reported in the text compare to a model with a seasonal AR(1) term with an additive MA(1) term?

c. Obtain the ACF for the growth rate of the M2NSA series. What type of model is suggested by the ACF?

d. Denote the seasonally differenced growth rate of M2NSA by $m2_t$. Estimate an AR(1) model with a seasonal MA term over the 1962:3 to 2014:4 period. You should obtain $m2_t = 0.5412 m2_{t-1} + \varepsilon_t - 0.8682\varepsilon_{t-4}$. Show that this model is preferable to (i) an AR(1) with a seasonal AR term, (ii) MA(1) with a seasonal AR term, and (iii) an MA(1) with a seasonal MA term.

e. Would you recommend including an MA term at lag 2 to remove any remaining serial correlation in the residuals?

16. The file labeled Y_BREAK.XLS contains the 150 observations of the series constructed as $y_t = 1 + 0.5y_{t-1} + (1 + 0.1y_{t-1})D_t + \varepsilon_t$ where D_t is a dummy variable equal to 0 for $t < 101$ and equal to 1.5 for $t \geq 101$.

a. Explain how this representation of the model allows the intercept to jump from 1 to 2.5 and the AR(1) coefficient to jump from 0.5 to 0.65.

b. Use the data to verify the results reported in the text.

c. Why do you think that the estimated intercept actually falls beginning with period 101?

d. Estimate the series as an AR(2) process. In what sense does the AR(2) model perform better than the AR(1) model estimated in part a?

e. Perform a recursive estimation of the AR(2) model and plot the CUSUMs. Is the AR(2) model adequate?