Chapter 9

Assembling the Components: U.S. Liquor Sales

Thus far we've focused on modeling trend, seasonals, and cycles one at a time. In Chapter 5, we introduced models and forecasts of trends and seasonality, respectively. Although cycles were likely present in the retail sales and housing starts series that we examined empirically, we simply ignored them. In Chapters 6 and 7 we introduced models and forecasts of cycles. We forecasted employment using autoregressive models. We didn't need trends or seasonals, because our employment series had no trend or seasonality.

In many forecasting situations, however, more than one component is needed to capture the dynamics in a series to be forecast – frequently they're *all* needed. Here we assemble our tools for forecasting trends, seasonals, and cycles; we use regression on a trend and calendar-effect dummies, and we capture cyclical dynamics by allowing for autoregressive effects in the regression disturbances, or by directly including lagged dependent variables in the regression.

9.1 Serially Correlated Disturbances

The full model is:

$$y_t = T_t(\theta) + \sum_{i=1}^{s} \gamma_i D_{it} + \varepsilon_t$$

$$\Phi(L)\varepsilon_t = v_t$$

$$\Phi(L) = 1 - \phi_1 L - ... - \phi_p L^p$$

$$v_t \sim WN(0, \sigma^2).$$

$T_t(\theta)$ is a trend, with underlying parameters θ. For example, linear trend has $\theta = \beta_1$ and

$$T_t(\theta) = \beta_1 TIME_t,$$

and quadratic trend has $\theta = (\beta_1, \beta_2)$ and

$$T_t(\theta) = \beta_1 TIME_t + \beta_2 TIME_t^2.$$

In addition to the trend, we include seasonal dummies.[1,2] The disturbances follow an $AR(p)$ process. In any particular application, of course, various trend effects, seasonal and other calendar effects, and autoregressive cyclical effects may not be needed and so could be dropped.[3] Finally, v_t is the underlying white noise shock that drives everything.

Now consider constructing an h-step-ahead point forecast at time T, $y_{T+h,T}$. At time $T+h$,

$$y_{T+h} = T_{T+h}(\theta) + \sum_{i=1}^{s} \gamma_i D_{i,T+h} + \varepsilon_{T+h}.$$

Projecting the right-hand side variables on what's known at time T (that is,

[1] Note that, because we include a full set of seasonal dummies, the trend does not contain an intercept, and we don't include an intercept in the regression.

[2] Holiday and trading-day dummies could could of course also be included if relevant.

[3] If the seasonal dummies were dropped, then we'd include an intercept in the regression.

the time-T information set, Ω_T), yields the point forecast

$$y_{T+h,T} = T_{T+h}(\theta) + \sum_{i=1}^{s} \gamma_i D_{i,T+h} + \varepsilon_{T+h,T}.$$

As with the pure trend and seasonal models discussed earlier, the trend and seasonal variables on the right-hand side are perfectly predictable. The only twist concerns the cyclical behavior that may be lurking in the disturbance term, future values of which don't necessarily project to zero, because the disturbance is no longer necessarily white noise. Instead, we construct $\varepsilon_{T+h,T}$ using the methods we developed for forecasting cycles.

As always, we make the point forecast operational by replacing unknown parameters with estimates, yielding

$$\hat{y}_{T+h,T} = T_{T+h}(\hat{\theta}) + \sum_{i=1}^{s\hat{\gamma}_i} D_{i,T+h} + \hat{\varepsilon}_{T+h,T}.$$

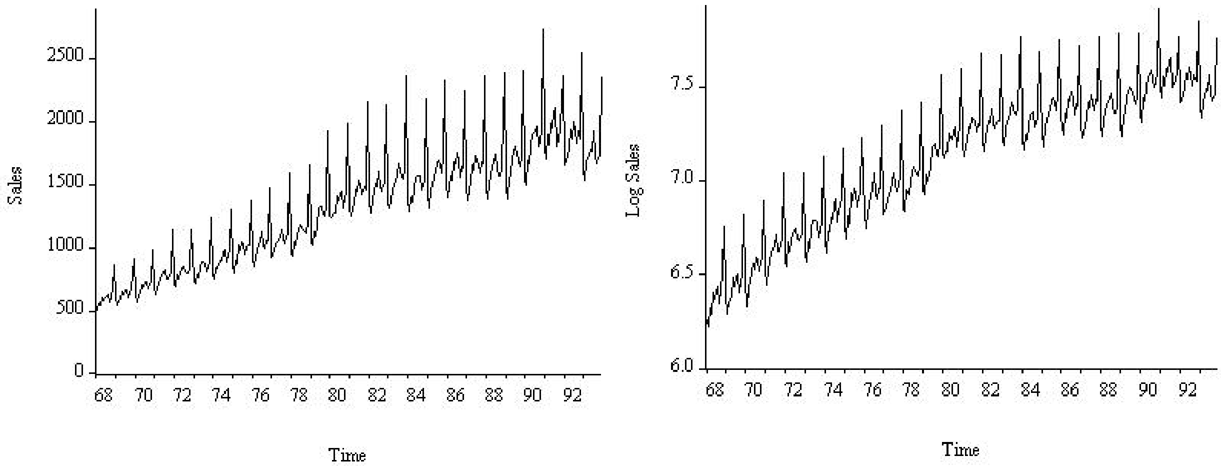
To construct $\hat{\varepsilon}_{T+h,T}$, in addition to replacing the parameters in the formula for $\varepsilon_{T+h,T}$ with estimates, we replace the unobservable disturbances, the ε_t's, with the observable residuals, the e_t's.

The complete h-step-ahead density forecast under normality is

$$N(\hat{y}_{T+h,T},\ \hat{\sigma}_h^2).$$

where $\hat{\sigma}_h^2$ is the operational estimate of the variance of the error in forecasting ε_{T+h}.

Once again, we don't actually have to *do* any of the computations just discussed; rather, the computer does them all for us. So let's get on with an application, now that we know what we're doing.



(a) Liquor Sales in Levels, 1968 - 1986

(b) Log Liquor Sales, 1968 - 1993

Figure 9.1: Liquor Sales

9.2 Lagged Dependent Variables

We use:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + T_t(\theta) + \sum_{i=1}^{s} \gamma_i D_{it} + \varepsilon_t$$

$$\varepsilon_t \sim WN(0, \sigma^2).$$

9.2.1 Case Study: Forecasting Liquor Sales with Deterministic Trends and Seasonals

We'll forecast monthly U.S. liquor sales. In Figure 9.1a, we show the history of liquor sales, 1968.01 - 1993.12. Notice its pronounced seasonality – sales skyrocket during the Christmas season. In Figure 9.1b we show log liquor sales; we take logs to stabilize the variance, which grows over time.[4] The variance of log liquor sales is more stable, and it's the series for which we'll build forecasting models.[5]

[4]The nature of the logarithmic transformation is such that it "compresses" an increasing variance. Make a graph of log(x) as a function of x, and you'll see why.

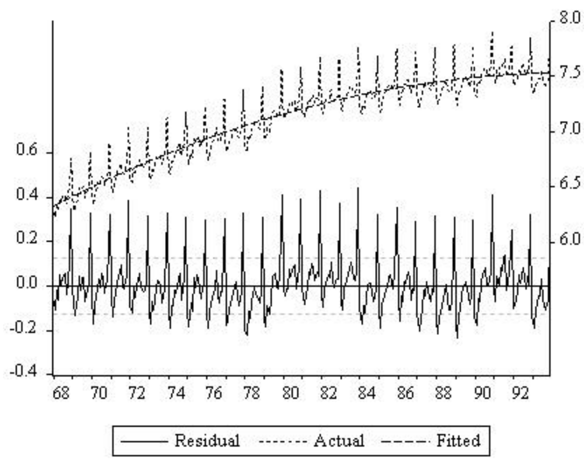
[5]From this point onward, for brevity we'll simply refer to "liquor sales," but remember that we've taken logs.

LS // Dependent Variable is LSALES
Sample: 1968:01 1993:12
Included observations: 312

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.237356	0.024496	254.6267	0.0000
TIME	0.007690	0.000336	22.91552	0.0000
TIME2	-1.14E-05	9.74E-07	-11.72695	0.0000

R-squared	0.892394	Mean dependent var	7.112383
Adjusted R-squared	0.891698	S.D. dependent var	0.379308
S.E. of regression	0.124828	Akaike info criterion	-4.152073
Sum squared resid	4.814823	Schwarz criterion	-4.116083
Log likelihood	208.0146	F-statistic	1281.296
Durbin-Watson stat	1.752858	Prob(F-statistic)	0.000000

(a) Liquor Sales, Quadratic Trend Regression



(b) Liquor Sales, Quadratic Trend Regression - Residual Plot

Figure 9.2: Liquor Sales: Quadratic Trend Model

Liquor sales dynamics also feature prominent trend and cyclical effects. Liquor sales trend upward, and the trend appears nonlinear in spite of the fact that we're working in logs. To handle the nonlinear trend, we adopt a quadratic trend model (in logs). The estimation results are in Table 9.2a. The residual plot (Figure 9.2b) shows that the fitted trend increases at a decreasing rate; both the linear and quadratic terms are highly significant. The adjusted R^2 is 89%, reflecting the fact that trend is responsible for a large part of the variation in liquor sales. The standard error of the regression is .125; it's an estimate of the standard deviation of the error we'd expect to make in forecasting liquor sales if we accounted for trend but ignored

	Acorr.	P. Acorr.	Std. Error	Ljung-Box	p-value
1	0.117	0.117	.056	4.3158	0.038
2	-0.149	-0.165	.056	11.365	0.003
3	-0.106	-0.069	.056	14.943	0.002
4	-0.014	-0.017	.056	15.007	0.005
5	0.142	0.125	.056	21.449	0.001
6	0.041	-0.004	.056	21.979	0.001
7	0.134	0.175	.056	27.708	0.000
8	-0.029	-0.046	.056	27.975	0.000
9	-0.136	-0.080	.056	33.944	0.000
10	-0.205	-0.206	.056	47.611	0.000
11	0.056	0.080	.056	48.632	0.000
12	0.888	0.879	.056	306.26	0.000
13	0.055	-0.507	.056	307.25	0.000
14	-0.187	-0.159	.056	318.79	0.000
15	-0.159	-0.144	.056	327.17	0.000
16	-0.059	-0.002	.056	328.32	0.000
17	0.091	-0.118	.056	331.05	0.000
18	-0.010	-0.055	.056	331.08	0.000
19	0.086	-0.032	.056	333.57	0.000
20	-0.066	0.028	.056	335.03	0.000
21	-0.170	0.044	.056	344.71	0.000
22	-0.231	0.180	.056	362.74	0.000
23	0.028	0.016	.056	363.00	0.000
24	0.811	-0.014	.056	586.50	0.000
25	0.013	-0.128	.056	586.56	0.000
26	-0.221	-0.136	.056	603.26	0.000
27	-0.196	-0.017	.056	616.51	0.000
28	-0.092	-0.079	.056	619.42	0.000
29	0.045	-0.094	.056	620.13	0.000
30	-0.043	0.045	.056	620.77	0.000
31	0.057	0.041	.056	621.89	0.000
32	-0.095	-0.002	.056	625.07	0.000
33	-0.195	0.026	.056	638.38	0.000
34	-0.240	0.088	.056	658.74	0.000
35	0.006	-0.089	.056	658.75	0.000
36	0.765	0.076	.056	866.34	0.000

Figure 9.3: Liquor Sales, Quadratic Trend - Residual Correlogram

seasonality and serial correlation. The Durbin-Watson statistic provides no evidence against the hypothesis that the regression disturbance is white noise.

The residual plot, however, shows obvious residual seasonality. The Durbin-Watson statistic missed it, evidently because it's not designed to have power against seasonal dynamics.[6] The residual plot also suggests that there may be a cycle in the residual, although it's hard to tell (hard for the Durbin-Watson statistic as well), because the pervasive seasonality swamps the picture and makes it hard to infer much of anything.

The residual correlogram (Table 9.3) and its graph (Figure 9.4) confirm the importance of the neglected seasonality. The residual sample autocor-

[6] Recall that the Durbin-Watson test is designed to detect simple AR(1) dynamics. It also has the ability to detect other sorts of dynamics, but evidently not those relevant to the present application, which are very different from a simple AR(1).

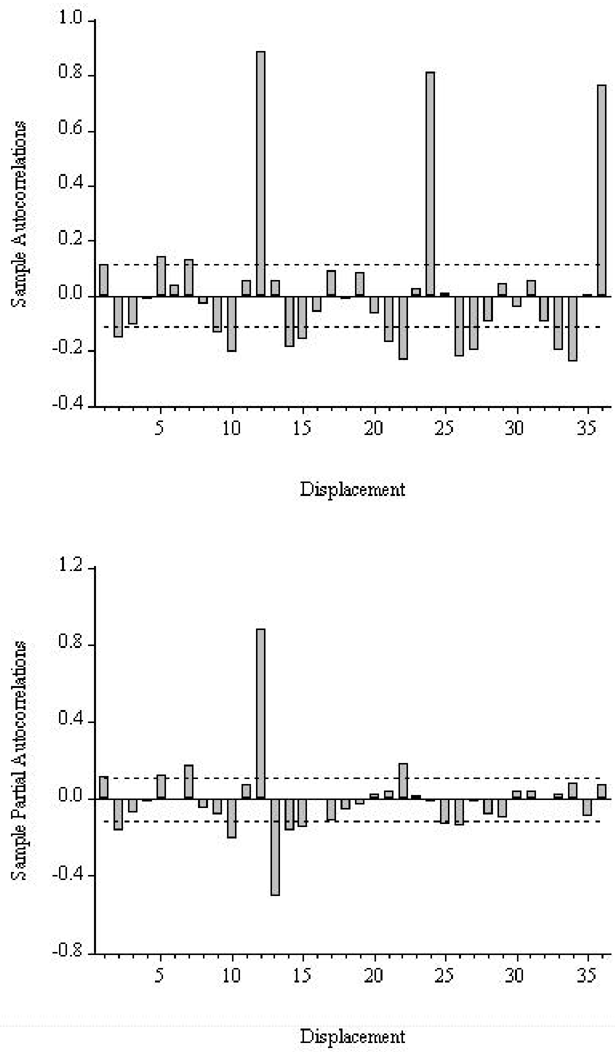


Figure 9.4: Liquor Sales, Quadratic Trend Regression - Residual Sample Autocorrelation

relation function has large spikes, far exceeding the Bartlett bands, at the seasonal displacements, 12, 24, and 36. It indicates some cyclical dynamics as well; apart from the seasonal spikes, the residual sample autocorrelation and partial autocorrelation functions oscillate, and the Ljung-Box statistic rejects the white noise null hypothesis even at very small, non-seasonal, displacements.

In Table 9.5a we show the results of regression on quadratic trend and a full set of seasonal dummies. The quadratic trend remains highly significant. The adjusted R^2 rises to 99%, and the standard error of the regression falls to .046, which is an estimate of the standard deviation of the forecast error we expect to make if we account for trend and seasonality but ignore serial correlation. The Durbin-Watson statistic, however, has greater ability to detect serial correlation now that the residual seasonality has been accounted for, and it sounds a loud alarm.

The residual plot of Figure 9.5b shows no seasonality, as that's now picked up by the model, but it confirms the Durbin-Watson's warning of serial correlation. The residuals are highly persistent, and hence predictable. We show the residual correlogram in tabular and graphical form in Table 9.6 and Figure 9.7. The residual sample autocorrelations oscillate and decay slowly, and they exceed the Bartlett standard errors throughout. The Ljung-Box test strongly rejects the white noise null at all displacements. Finally, the residual sample partial autocorrelations cut off at displacement 3. All of this suggests that an AR(3) would provide a good approximation to the disturbance's Wold representation.

In Table 9.8a, then, we report the results of estimating a liquor sales model with quadratic trend, seasonal dummies, and AR(3) disturbances. The R^2 is now 100%, and the Durbin-Watson is fine. One inverse root of the AR(3) disturbance process is estimated to be real and close to the unit circle (.95), and the other two inverse roots are a complex conjugate pair farther from the unit circle. The standard error of this regression is an estimate of the standard deviation of the forecast error we'd expect to make after modeling the residual serial correlation, as we've now done; that is, it's an estimate of the standard deviation of v.[7]

We show the residual plot in Figure 9.8b and the residual correlogram in Table 9.9 and Figure fig: liquor sales quadratic seasonal dummies and ar(3) residual sample autocorrelation. The residual plot reveals no patterns; instead, the residuals look like white noise, as they should. The residual

[7] Recall that v is the innovation that drives the ARMA process for the regression disturbance, ε. It's a very small .027, roughly half that obtained when we ignored serial correlation.

sample autocorrelations and partial autocorrelations display no patterns and are mostly inside the Bartlett bands. The Ljung-Box statistics also look good for small and moderate displacements, although their p-values decrease for longer displacements.

All things considered, the quadratic trend, seasonal dummy, AR(3) specification seems tentatively adequate. We also perform a number of additional checks. In Figure 9.11, we show a histogram and normality test applied to the residuals. The histogram looks symmetric, as confirmed by the skewness near zero. The residual kurtosis is a bit higher then three and causes Jarque-Bera test to reject the normality hypothesis with a p-value of .02, but the residuals nevertheless appear to be fairly well approximated by a normal distribution, even if they may have slightly fatter tails.

Now we use the estimated model to produce forecasts. In Figure 9.12 we show the history of liquor sales and a 12-month-ahead extrapolation forecast for 1994.[8] To aid visual interpretation, we show only two years of history. The forecast looks reasonable. It's visually apparent that the model has done a good job of picking up the seasonal pattern, which dominates the local behavior of the series. In Figure 9.13, we show the history, the forecast, and the 1994 realization. The forecast was very good!

In Figure 9.14 we show four years of history together with a 60-month-ahead (five year) extrapolation forecast, to provide a better feel for the dynamics in the forecast. The figure also makes clear the trend forecast is slightly *downward*. To put the long-horizon forecast in historical context, we show in Figure 13 the 60-month-ahead forecast together with the complete history. Finally, in Figure 14, we show the history and point forecast of the level of liquor sales (as opposed to log liquor sales), which we obtain by exponentiating the forecast of log liquor sales.[9]

[8] We show the point forecast together with 95% intervals.

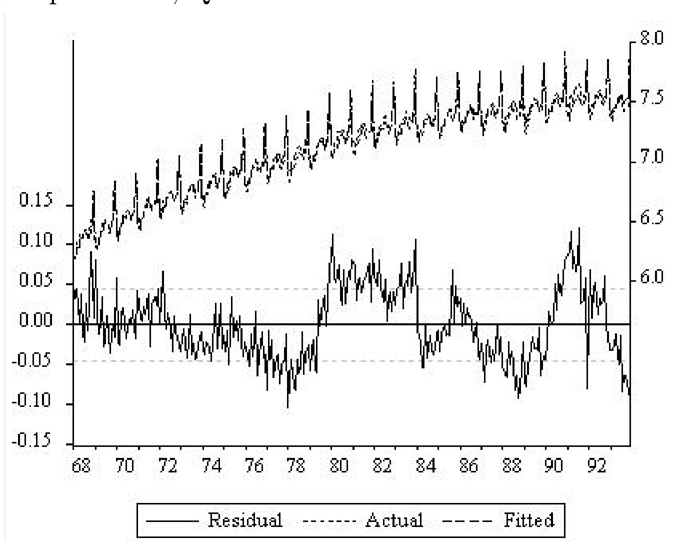
[9] Recall that exponentiating "undoes" a natural logarithm.

LS // Dependent Variable is LSALES
Sample: 1968:01 1993:12
Included observations: 312

Variable	Coefficient	Std. Error	t-Statistic	Prob.
TIME	0.007656	0.000123	62.35882	0.0000
TIME2	-1.14E-05	3.56E-07	-32.06823	0.0000
D1	6.147456	0.012340	498.1699	0.0000
D2	6.088653	0.012353	492.8890	0.0000
D3	6.174127	0.012366	499.3008	0.0000
D4	6.175220	0.012378	498.8970	0.0000
D5	6.246086	0.012390	504.1398	0.0000
D6	6.250387	0.012401	504.0194	0.0000
D7	6.295979	0.012412	507.2402	0.0000
D8	6.268043	0.012423	504.5509	0.0000
D9	6.203832	0.012433	498.9630	0.0000
D10	6.229197	0.012444	500.5968	0.0000
D11	6.259770	0.012453	502.6602	0.0000
D12	6.580068	0.012463	527.9819	0.0000

R-squared	0.986111	Mean dependent var	7.112383
Adjusted R-squared	0.985505	S.D. dependent var	0.379308
S.E. of regression	0.045666	Akaike info criterion	-6.128963
Sum squared resid	0.621448	Schwarz criterion	-5.961008
Log likelihood	527.4094	F-statistic	1627.567
Durbin-Watson stat	0.586187	Prob(F-statistic)	0.000000

(a) Liquor Sales, Quadratic Trend with Seasonal Dummies



(b) Liquor Sales, Quadratic Trend with Seasonal Dummies - Residual Plot

Figure 9.5: Liquor Sales - Trend and Seasonal Model

	Acorr.	P. Acorr.	Std. Error	Ljung-Box	p-value
1	0.700	0.700	.056	154.34	0.000
2	0.686	0.383	.056	302.86	0.000
3	0.725	0.369	.056	469.36	0.000
4	0.569	-0.141	.056	572.36	0.000
5	0.569	0.017	.056	675.58	0.000
6	0.577	0.093	.056	782.19	0.000
7	0.460	-0.078	.056	850.06	0.000
8	0.480	0.043	.056	924.38	0.000
9	0.466	0.030	.056	994.46	0.000
10	0.327	-0.188	.056	1029.1	0.000
11	0.364	0.019	.056	1072.1	0.000
12	0.355	0.089	.056	1113.3	0.000
13	0.225	-0.119	.056	1129.9	0.000
14	0.291	0.065	.056	1157.8	0.000
15	0.211	-0.119	.056	1172.4	0.000
16	0.138	-0.031	.056	1178.7	0.000
17	0.195	0.053	.056	1191.4	0.000
18	0.114	-0.027	.056	1195.7	0.000
19	0.055	-0.063	.056	1196.7	0.000
20	0.134	0.089	.056	1202.7	0.000
21	0.062	0.018	.056	1204.0	0.000
22	-0.006	-0.115	.056	1204.0	0.000
23	0.084	0.086	.056	1206.4	0.000
24	-0.039	-0.124	.056	1206.9	0.000
25	-0.063	-0.055	.056	1208.3	0.000
26	-0.016	-0.022	.056	1208.4	0.000
27	-0.143	-0.075	.056	1215.4	0.000
28	-0.135	-0.047	.056	1221.7	0.000
29	-0.124	-0.048	.056	1227.0	0.000
30	-0.189	0.086	.056	1239.5	0.000
31	-0.178	-0.017	.056	1250.5	0.000
32	-0.139	0.073	.056	1257.3	0.000
33	-0.226	-0.049	.056	1275.2	0.000
34	-0.155	0.097	.056	1283.7	0.000
35	-0.142	0.008	.056	1290.8	0.000
36	-0.242	-0.074	.056	1311.6	0.000

Figure 9.6: Liquor Sales, Quadratic Trend with Seasonal Dummies - Residual Correlogram

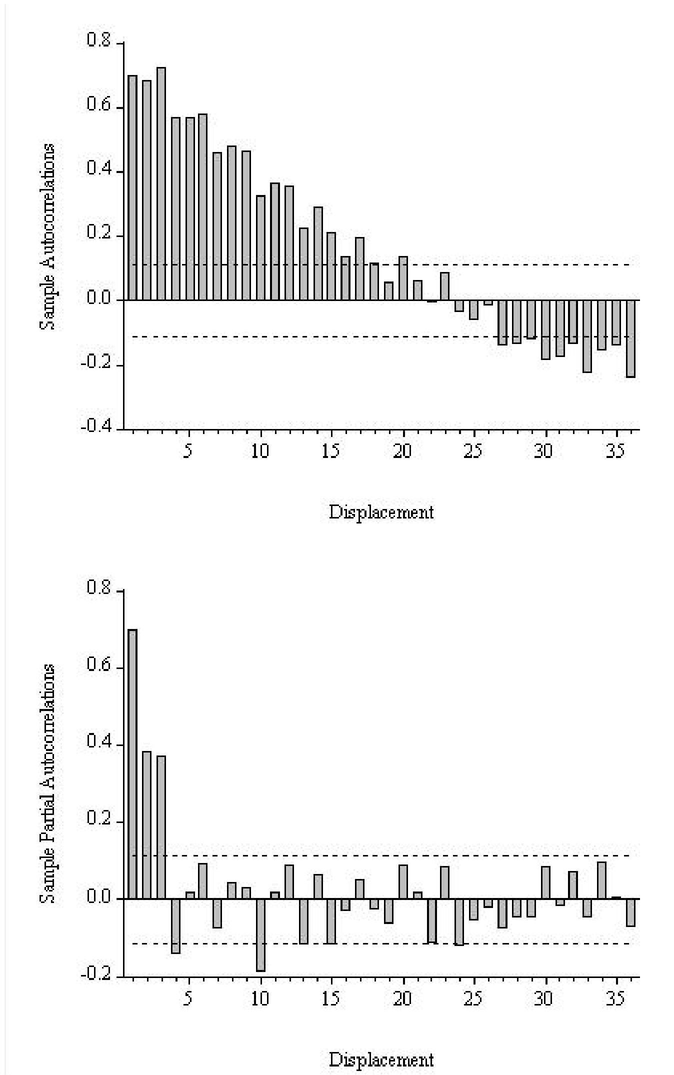


Figure 9.7: Liquor Sales, Quadratic Trend with Seasonal Dummies - Residual Sample Autocorrelation

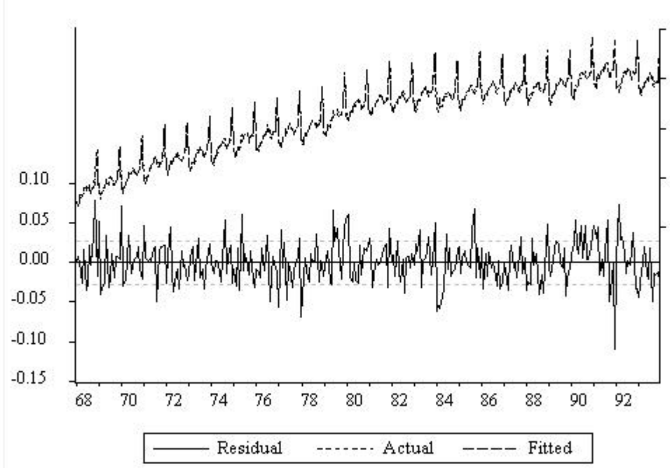
LS // Dependent Variable is LSALES
Sample: 1968:01 1993:12
Included observations: 312
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
TIME	0.008606	0.000981	8.768212	0.0000
TIME2	-1.41E-05	2.53E-06	-5.556103	0.0000
D1	6.073054	0.083922	72.36584	0.0000
D2	6.013822	0.083942	71.64254	0.0000
D3	6.099208	0.083947	72.65524	0.0000
D4	6.101522	0.083934	72.69393	0.0000
D5	6.172528	0.083946	73.52962	0.0000
D6	6.177129	0.083947	73.58364	0.0000
D7	6.223323	0.083939	74.14071	0.0000
D8	6.195681	0.083943	73.80857	0.0000
D9	6.131818	0.083940	73.04993	0.0000
D10	6.157592	0.083934	73.36197	0.0000
D11	6.188480	0.083932	73.73176	0.0000
D12	6.509106	0.083928	77.55624	0.0000
AR(1)	0.268805	0.052909	5.080488	0.0000
AR(2)	0.239688	0.053697	4.463723	0.0000
AR(3)	0.395880	0.053109	7.454150	0.0000

R-squared	0.995069	Mean dependent var	7.112383
Adjusted R-squared	0.994802	S.D. dependent var	0.379308
S.E. of regression	0.027347	Akaike info criterion	-7.145319
Sum squared resid	0.220625	Schwarz criterion	-6.941373
Log likelihood	688.9610	F-statistic	3720.875
Durbin-Watson stat	1.886119	Prob(F-statistic)	0.000000

Inverted AR Roots .95 -.34+.55i -.34 -.55i

(a) Liquor Sales, Quadratic Trend with Seasonal Dummies and AR(3)



(b) Liquor Sales, Quadratic Trend with Seasonal Dummies and AR(3) Disturbances - Residual Plot

Figure 9.8: Liquor Sales - Trend, Seasonal, and AR(3) Model

	Acorr.	P. Acorr.	Std. Error	Ljung-Box	p-value
1	0.056	0.056	.056	0.9779	0.323
2	0.037	0.034	.056	1.4194	0.492
3	0.024	0.020	.056	1.6032	0.659
4	-0.084	-0.088	.056	3.8256	0.430
5	-0.007	0.001	.056	3.8415	0.572
6	0.065	0.072	.056	5.1985	0.519
7	-0.041	-0.044	.056	5.7288	0.572
8	0.069	0.063	.056	7.2828	0.506
9	0.080	0.074	.056	9.3527	0.405
10	-0.163	-0.169	.056	18.019	0.055
11	-0.009	-0.005	.056	18.045	0.081
12	0.145	0.175	.056	24.938	0.015
13	-0.074	-0.078	.056	26.750	0.013
14	0.149	0.113	.056	34.034	0.002
15	-0.039	-0.060	.056	34.532	0.003
16	-0.089	-0.058	.056	37.126	0.002
17	0.058	0.048	.056	38.262	0.002
18	-0.062	-0.050	.056	39.556	0.002
19	-0.110	-0.074	.056	43.604	0.001
20	0.100	0.056	.056	46.935	0.001
21	0.039	0.042	.056	47.440	0.001
22	-0.122	-0.114	.056	52.501	0.000
23	0.146	0.130	.056	59.729	0.000
24	-0.072	-0.040	.056	61.487	0.000
25	0.006	0.017	.056	61.500	0.000
26	0.148	0.082	.056	69.024	0.000
27	-0.109	-0.067	.056	73.145	0.000
28	-0.029	-0.045	.056	73.436	0.000
29	-0.046	-0.100	.056	74.153	0.000
30	-0.084	0.020	.056	76.620	0.000
31	-0.095	-0.101	.056	79.793	0.000
32	0.051	0.012	.056	80.710	0.000
33	-0.114	-0.061	.056	85.266	0.000
34	0.024	0.002	.056	85.468	0.000
35	0.043	-0.010	.056	86.116	0.000
36	-0.229	-0.140	.056	104.75	0.000

Figure 9.9: Liquor Sales, Quadratic Trend with Seasonal Dummies and AR(3) Disturbances - Residual Correlogram

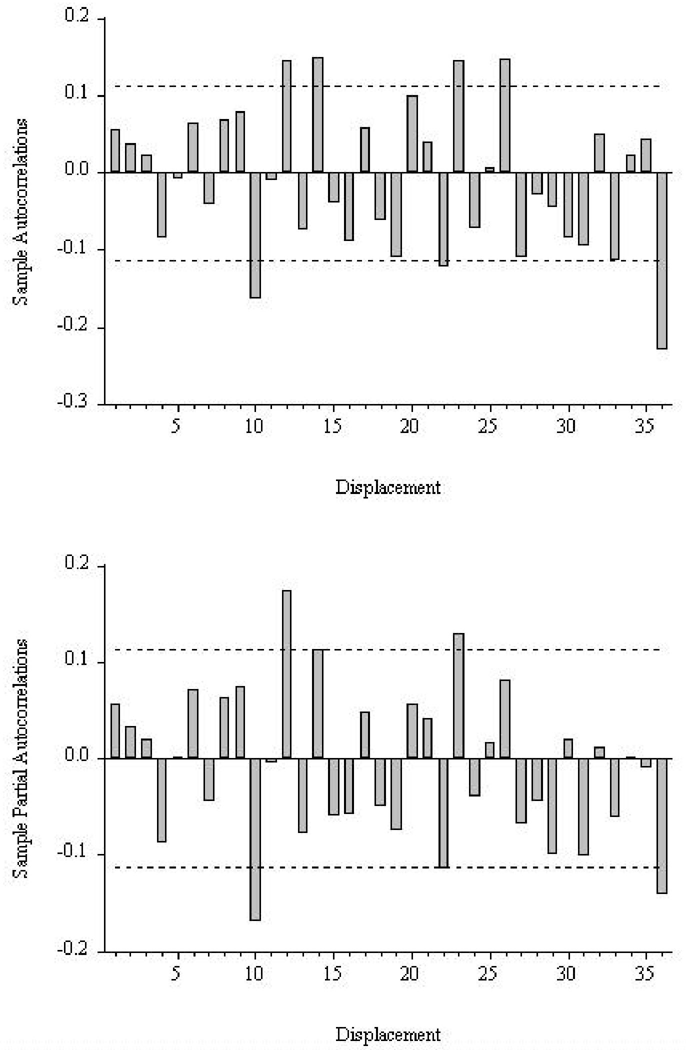


Figure 9.10: Liquor Sales, Quadratic Trend with Seasonal Dummies and AR(3) Disturbances - Residual Sample Autocorrelation

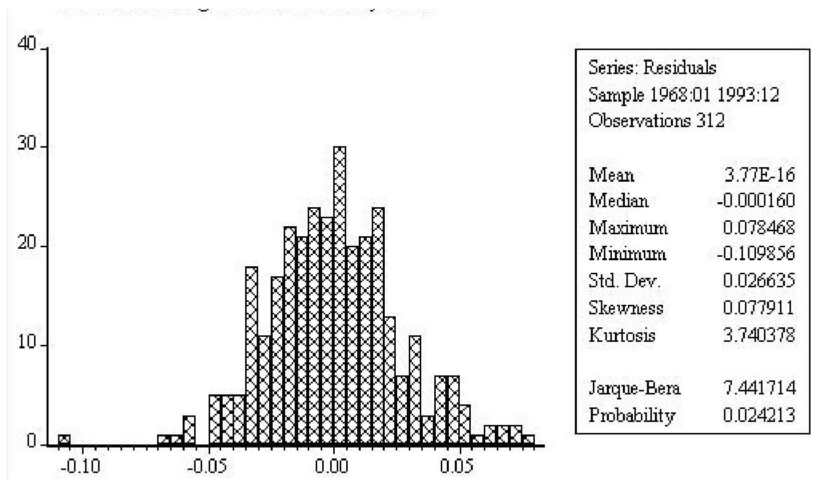


Figure 9.11: Liquor Sales, Quadratic Trend with Seasonal Dummies and AR(3) Disturbances - Residual Histogram and Normality test

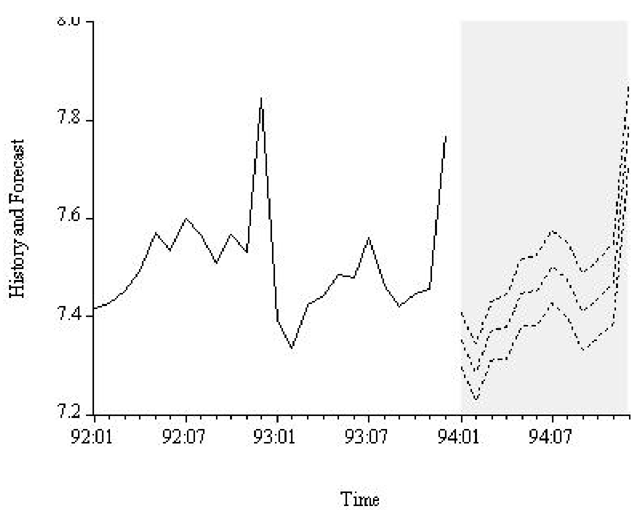


Figure 9.12: Liquor Sales: History and 12-Month-Ahead Forecast

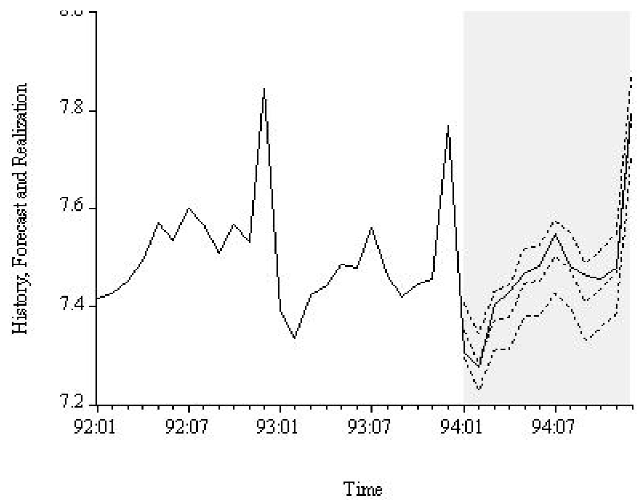


Figure 9.13: Liquor Sales: History, 12-Month-Ahead Forecast, and Realization

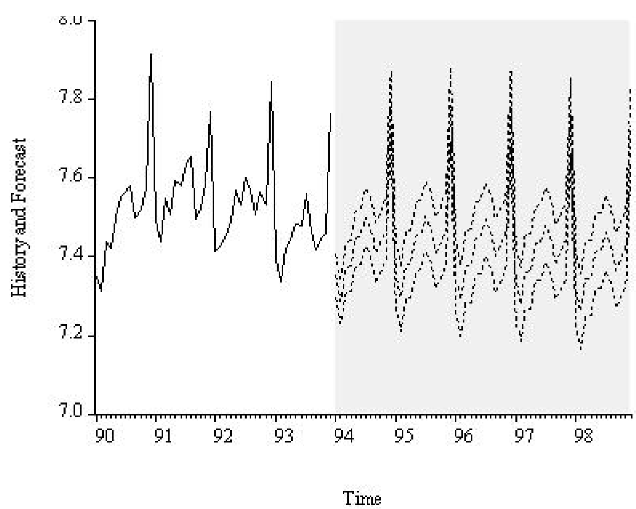


Figure 9.14: Liquor Sales: History and Four-Year-Ahead Forecast

9.3 Exercises, Problems and Complements

1. Serially correlated disturbances vs. lagged dependent variables. Estimate the quadratic trend model for log liquor sales with seasonal dummies and three lags of the dependent variable included directly. Discuss your results and compare them to those we obtained when we instead allowed for AR(3) disturbances in the regression. Which model is selected by AIC and SIC?

2. Assessing the adequacy of the liquor sales forecasting model deterministic trend specification. Critique the liquor sales forecasting model that we adopted (log liquor sales with quadratic trend, seasonal dummies, and AR(3) disturbances).[10]

   a. If the trend is not a good approximation to the actual trend in the series, would it greatly affect short-run forecasts? Long-run forecasts?

   b. How might you fit and assess the adequacy of a broken linear trend? How might you decide on the location of the break point?

3. Improving non-trend aspects of the liquor sales forecasting model.

   a. Recall our argument that best practice requires using a χ^2_{m-k} distribution rather than a χ^2_m distribution to assess the significance of Q-statistics for model residuals, where m is the number of autocorrelations included in the Q statistic and k is the number of parameters estimated. In several places in this chapter, we failed to heed this advice when evaluating the liquor sales model. If we were instead to compare the residual Q-statistic p-values to a χ^2_{m-k} distribution, how, if at all, would our assessment of the model's adequacy change?

   b. Return to the log-quadratic trend model with seasonal dummies, allow for $ARMA(p, q)$ disturbances, and do a systematic selection of p and

[10] I thank Ron Michener, University of Virginia, for suggesting parts d and f.

q using AIC and SIC. Do AIC and SIC select the same model? If not, which do you prefer? If your preferred disturbance model differs from the $AR(3)$ that we used, replicate the analysis in the text using your preferred model, and discuss your results.

c. Discuss and evaluate another possible model improvement: inclusion of an additional dummy variable indicating the number of Fridays and/or Saturdays in the month. Does this model have lower AIC or SIC than the final model used in the text? Do you prefer it to the one in the text? Why or why not?

9.4 Notes