ECO 5375-701 | Prof. Tom Fomby
Eco and Bus Forecasting | Fall 2016

EXERCISE 2
KEY

Purpose: To learn how to use the DTDS model to **test for the presence or absence of seasonality in time series data** and to estimate a DTDS model for the Plano Sales Tax Revenue data. We are also going to identify the "strong" versus the "weak" seasons during the year as it relates to this data. You are to hand in this exercise in class on Thursday, September 22.

You are to use the SAS program Plano_Test_Seasonality.sas to complete this exercise.

a) Using the first Proc Reg output, fill in the blanks in the below Durbin-Watson table:

The REG Procedure
Model: MODEL1
Dependent Variable: rev

Durbin-Watson D	1.532
Pr < DW	**_0.0006_**
Pr > DW	**_0.9994_**
Number of Observations	190
1st Order Autocorrelation	0.225

Note: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value for testing negative autocorrelation

Do the **ordinary least squares residuals** of our DTDS regression model seem to have autocorrelated errors? (**Yes** / No) Explain your answer.

Answer: It is the Pr<DW outcome that we are interested in. This represents the null hypothesis of no autocorrelation versus the alternative hypothesis of autocorrelated errors (a one-tailed test). Since the one-tail probability value of the DW test statistic is 0.0006 < 0.05 we reject the null hypothesis of no autocorrelation and accept the alternative hypothesis of positive autocorrelation in the errors of the least squares model. This result then indicates that we should conduct all of our subsequent hypothesis tests using Generalized Least Squares which corrects for autocorrelated errors (via Proc Autoreg).

b) Using the **first** Proc Autoreg output, fill in the blanks below:

Coefficient on t^2 variable = __**-31.3134**_____________.
T-statistic on t^2 variable = ___**-1.49**___________.
Therefore, we need to (include / **exclude**) the t^2 variable from further consideration.

c) Using the **second** Proc Autoreg output (the one without the t^2 term), fill in the blanks below.

These are the Durbin-Watson test statistics based on the OLS residuals. Fill in the blanks below.

Durbin-Watson Statistics

Order	DW	Pr < DW	Pr > DW
1	**1.3805**	<.0001	1.0000
2	1.5470	**0.0013**	0.9987
3	0.6783	<.0001	1.0000
4	1.4171	<.0001	0.9999

These are autocorrelation coefficients and their statistics produced by the second Proc Autoreg procedure. Fill in the blanks below.

AR1 Coefficient = ___**-0.2048**______________.
T-statistic on AR1 Coefficient = ___**-3.46**____________.
Therefore, this AR coefficient (**is** / is not) statistically significant.
AR3 Coefficient = ____**-0.5339**_______________.
T-statistic on AR3 Coefficient = ____**-9.11**___________.
Therefore, this AR coefficient (**is** / is not) statistically significant.
AR5 Coefficient = ______**-0.2524**_____________.
T-statistic on AR5 Coefficient = ____**-3.56**___________.
Therefore, this AR coefficient (**is** / is not) statistically significant.
AR8 Coefficient = _______**0.2119**____________.
T-statistic on AR8 Coefficient = ____**3.01**___________.
Therefore, this AR coefficient (**is** / is not) statistically significant.
AR10 Coefficient = _____**0.1852**_____________.
T-statistic on AR10 Coefficient = ____**2.98**___________.
Therefore, this AR coefficient (**is** / is not) statistically significant.
AR12 Coefficient = ______**-0.2540**_____________.
T-statistic on AR12 Coefficient = ____**-4.07**___________.
Therefore, this AR coefficient (**is** / is not) statistically significant.

The below Durbin-Watson table is based on the **Generalized least squares Residuals.** Fill in the blanks.

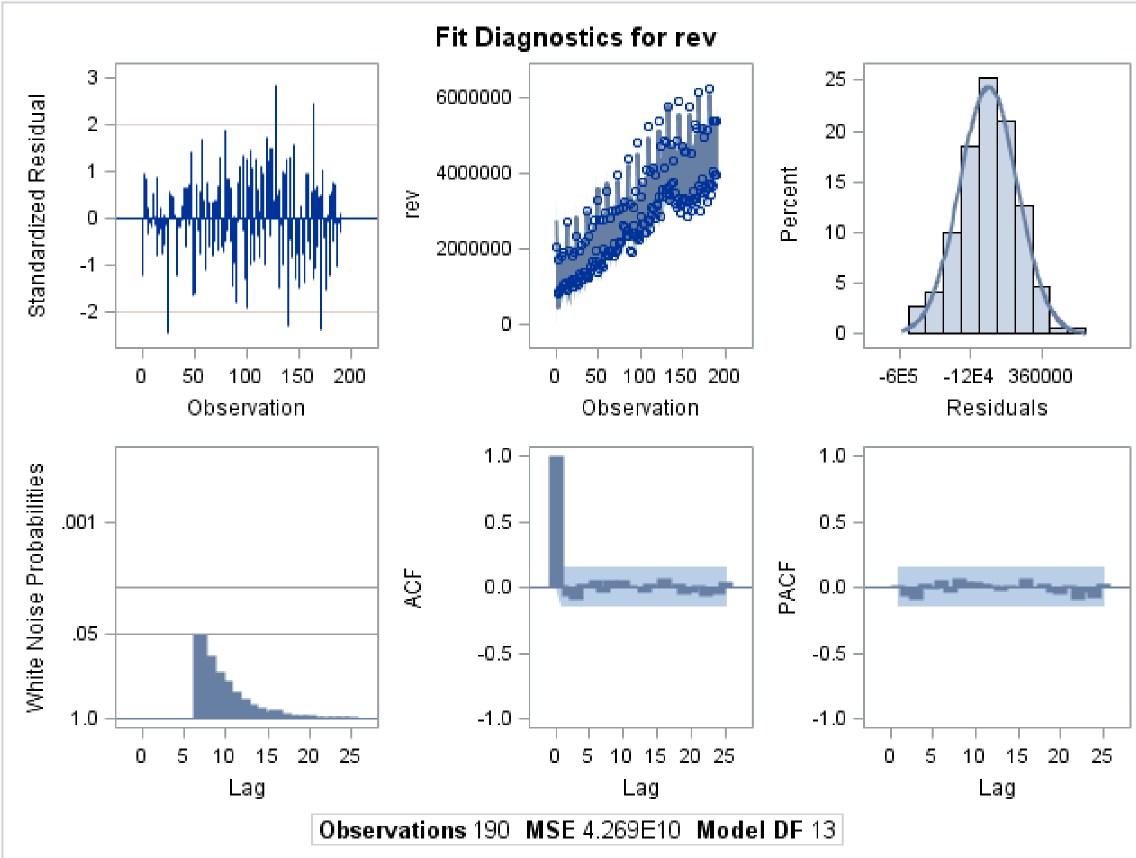
Durbin-Watson Statistics

Order	DW	Pr < DW	Pr > DW
1	1.9442	_**0.3406**_	0.6594
2	_**2.0388**_	0.6132	0.3868
3	2.1030	0.7982	_**0.2018**_
4	1.8571	0.2159	0.7841

NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value for testing negative autocorrelation.

True or False: Given the results in parts a) and c) above, it is obvious that we should be basing our statistical conclusions on **generalized least squares statistics** (Proc Autoreg) rather than **ordinary least squares statistics** (Proc Reg). Review the results you see in the above two Durbin-Watson tables and the graphs below. Explain how they suggest that the Proc Autoreg fit of the data is producing generalized least squares residuals that are white noise.

Answer: The first Durbin-Watson table involves OLS residuals and they appear to be autocorrelated because all of the test statistics have p-values less than 0.05. In contrast, the second Durbin-Watson table involves the generalized least squares residuals generated by the Proc Autoreg procedure. The results there produce p-values that are greater than 0.05 indicating that the generalized least squares residuals are uncorrelated. In addition, in the below graphs, the generalized least squares residuals have ACF and PACF functions that have no significant spikes and the Box-Pierce-Ljung statistics at the various lags are all statistically insignificant. Thus, given the white noise errors of the fitted Proc Autoreg model, we can judge our generalized least squares modeling to be complete.



d) Using the **second** Proc Autoreg output we can see that individual t-statistics indicate that the seasonal dummies __**for February, April, May, August, and November** _ are statistically significant at the 5% level.

In the below space write out in **conventional form** the model we have used Generalized least squares to fit the data including the coefficient estimates and the t-statistics in parentheses below the estimates. Be sure to write out the AR(1,3,5,8,10,12) model for the errors as well.

$$Y = 684937 + 18217t + 2043079D2 - 71393D3 - 176454D4 + 1118575D5 + 81750D6$$

(12.28) (12.28) (13.26) (-0.46) (-2.29) (7.28) (0.55)

$$+ 9397D7 + 1121733D8 + 54102D9 + 43245D10 + 1173353D11 - 82884D12 + E(t)$$

(0.12) (7.49) (0.36) (0.56) (7.48) (-0.54)

$$E(t) = 0.2048E(t-1) + 0.5339E(t-3) + 0.2524E(t-5) - 0.2119E(t-8) - 0.1852E(t-10)$$

(3.46) (9.11) (3.56) (-3.01) (-2.098)

$$+ 0.2540E(t-12) + a(t)$$

(4.07)

Note: In SAS Notation -ARi = rho(i) where rho(i) is the autoregressive coefficient that I discussed in class. For example, in the output AR1 = -0.2048 whereas, in the notation I used in class, the rho(1) coefficient is 0.2048. See the above.

e) A comprehensive test of the significance of seasonality in the Plano Tax Revenue data is based on an overall test of the following hypotheses using Generalized least squares (Proc Autoreg):

$H_0 : \gamma_2 = \gamma_3 = \cdots = \gamma_{12} = 0$ (No seasonality)

H_1 : At least one of the above γ's is not equal to zero (Seasonality is present)

Fill in the following F-table for this test:

Test 'Joint Test for Seasonality'

Source	**DF**	**Mean Square**	**F Value**	**Pr > F**
Numerator	11	1.5376516E12	**_36.02_**	<.0001
Denominator	171	42685238643		

Therefore, we should (accept / **reject**) the null hypothesis of no seasonality and should (**accept** / reject) the alternative hypothesis of seasonality in the Plano Tax Revenue data because **__the p-value associated with this joint test of the significance of the seasonal dummy coefficients is less than 0.05 and we accept the alternative hypothesis of seasonality in the data.**

f) According to the "Standardized seasonal effects by month" table generated by the Plano_Test_Seasonality.sas program the "strong" months in terms of the collection of tax revenue by Plano are **____February, May, August, and November_____**.
The "weak" months are **_January, March, April, June, July, September, October, and December__**.
The **strongest** month is **____February_____**. The **weakest** month is **___April_____**.