Name Mr. Key

ID# 7777777

ECO 5375
Eco. & Bus. Forecasting

Prof. Tom Fomby
Spring 2013

MID-TERM EXAM

Instructions: Fill in your name and student ID above. You have 1 hour and 20 minutes to complete this exam. This exam is worth a total of 66 points. The points for the separate questions are broken out as follows:

Q1 = 2 pts.
Q2 = 2 pts.
Q3 = 2 pts.
Q4 = 2 pts.
Q5 = 2 pts.
Q6 = 2 pts.
Q7 = 3 pts.
Q8 = 3 pts.
Q9 = 2 pts.
Q10 = 3 pts.
Q11 = 3 pts.
Q12 = 4 pts.
Q13 = 3 pts.
Q14 = 2 pts.
Q15 = 2 pts.
Q16 = 2 pts.
Q17 = 2 pts.
Q18 = ~~3~~ pts. 2
Q19 = 2 pts.
Q20 = 2 pts.
Q21 = 2 pts.
Q22 = 2 pts.
Q23 = 2 pts.
Q24 = 2 pts.
Q25 = 2 pts.
Q26 = 3 pts.
Q27 = 3 pts.
Q28 = 3 pts.

MULTIPLE CHOICE AND SHORT ANSWER QUESTIONS:

1. One way that we know that "Economic and Business Forecasting" is a bona fide field in economics is because:

2

a. Most economics departments offer a course in the subject
b. Several professional journals publish articles on the subject
c. You can obtain a certificate in forecasting from the Institute for Business Forecasting
d. All of the above

2. Which of the following problems involve predicting an level of output to produce and then deciding on how much capital and labor to use?

2

a. Coal Miner's Daughter Problem
b. Inventory Manager's Problem
c. Plant Manager's Problem
d. Private Investor Problem

3. SAS programs have two basic steps. They are the ___DATA___ step and the ___PROC___ step.
(procedure)

2

Consider the following Deterministic Trend/Deterministic Season model **for quarterly data**:

$$\hat{y}_t = 20 + 2t + 3D_{t2} - 3D_{t3} + 4D_{t4}.$$

Answer the following four questions using this information.

2 4. If we had also added a seasonal dummy variable for the first quarter we would have fallen into the so-called ___Dummy Variable___ trap.

2 5. The trend line for all first quarter data across all years is given by the formula
___$y = 20 + 2t$___

2 6. The trend line for all third quarter data across all years is given by the formula
___$y = 20 - 3 + 2t = 17 + 2t$___

3 7. Using the above equation calculate the mean intercept across all of the 4 quarters:
It is $\bar{\alpha} =$ ___21___. Show your work below.
Intercepts: 20(Qtr.1) + 23(Qtr.2) + 17(Qtr.3) + 24(Qtr.4)/4 = $\frac{84}{4}$ = 21

3 8. Calculate the relative strength of the third quarter, $\gamma_3^* =$ ___$-\frac{4}{21}$___. Show your work below.
Therefore, the ~~second~~ third quarter is a (weak / strong) quarter. (Circle the correct alternative)
$\gamma_3^* = \frac{17-21}{21} = \frac{-4}{21}$

Here are some Deterministic Trend/Deterministic Seasonal (DTDS) Model questions:

2 9. True or False. To test for the presence or absence of curvature in the trend of a time series when using the DTDS model, we test for the significance of the t^2 term in the model. If it is statistically

significant, then there is curvature in the trend. If it is insignificant, then the trend is linear.

10. Suppose that you were given the following Durbin-Watson output produced by an **OLS estimation** of a DTDS model.

3

Durbin-Watson Statistics			
Order	DW	Pr < DW	Pr > DW
1	0.7148	<.0001	1.0000
2	1.2929	<.0001	0.9999
3	1.4242	0.0013	0.9987
4	1.3938	0.0009	0.9991

Briefly describe to me the implications of this output as it relates to testing, say, the curvature parameter of the DTDS model and other tests of hypotheses like the presence or absence of seasonality. This output indicates that there is positive autocorrelation in the errors of the DTDS model. Therefore all tests of hypotheses should proceed by using Generalized Least Squares (i.e. PROC AUTOREG) not Ordinary Least Squares (PROC REG).

11. Consider the following F-test output derived from a GLS estimation of a DTDS model as motivated by the output in question 10 above.

3

Test Season				
Source	DF	Mean Square	F Value	Pr > F
Numerator	11	34109	245.60	<.0001
Denominator	106	138.877520		

Given this information, explain to me your conclusion concerning the presence or absence of seasonality in the time series being investigated. Be sure and tell me the **null and alternative hypotheses** in this test and how you drew your conclusion. H_0: There is no seasonality in the data

H_1: there is seasonality in the data. As the F-statistic's p-value is less than the conventional $\alpha = 0.05$ we reject H_0 and accept H_1. There appears to be seasonality in the data.

12. Here is a table of standardized seasonal effects generated by a DTDS model:

Obs	sum	d1a	d2a	d3a	d4a	d5a	d6a	d7a	d8a
1	-2.7756E-16	-0.41046	1.43772	-0.44900	-0.56675	0.60094	-0.30948	-0.39709	0.60483

Obs	d9a	d10a	d11a	d12a
1	-0.33027	-0.37695	0.65101	-0.45451

(1) The "weak" months are 1, 3, 4, 6, 7, 9, 10, 12 Jan, Mar. Apr. Jun. July, Sept. Oct. Dec.

(1) The "weakest" of the weak months is April (month 4.) most negative

(1) The "strong" months are 2, 5, 8, 11 Feb. May Aug. Nov.

(1) The "strongest" of the strong months is Feb. (Month 2.) most positive

Here are some **Box-Jenkins** questions:

(3) 13. A time series is stationary if it has constant mean, constant variance, and constant covariance.

(2) 14. In the Box-Jenkins approach, a popular transformation for non-stationary is the first difference of the data.

(2) 15. True or False: **Slowly-turning** series are typically non-stationary and must be transformed to make them stationary.

function

(2) 16. Before the Dickey-Fuller test for unit roots, the way that analysts determined non-stationarity of time series was to use the autocorrelation function. Briefly describe to me how the autocorrelation function was used to distinguish between time series that are stationary versus times series that are not. If the autocorrelation is "quickly" damping then the data are stationary. In contrast, if ACF is "slowly" damping then the data are nonstationary and need to be transformed to stationarity.

17. If a time series y_t exhibits an exponential growth pattern, it can be transformed to stationarity by the transformation (circle the correct transformation)

(2)

a. Δy_t

b. $\Delta \log_e(y_t)$

c. $\exp(y_t)$

d. $\tan^{-1}(y_t)$

Consider the below four graphs of time series data and then answer the four questions that follow the graphs:

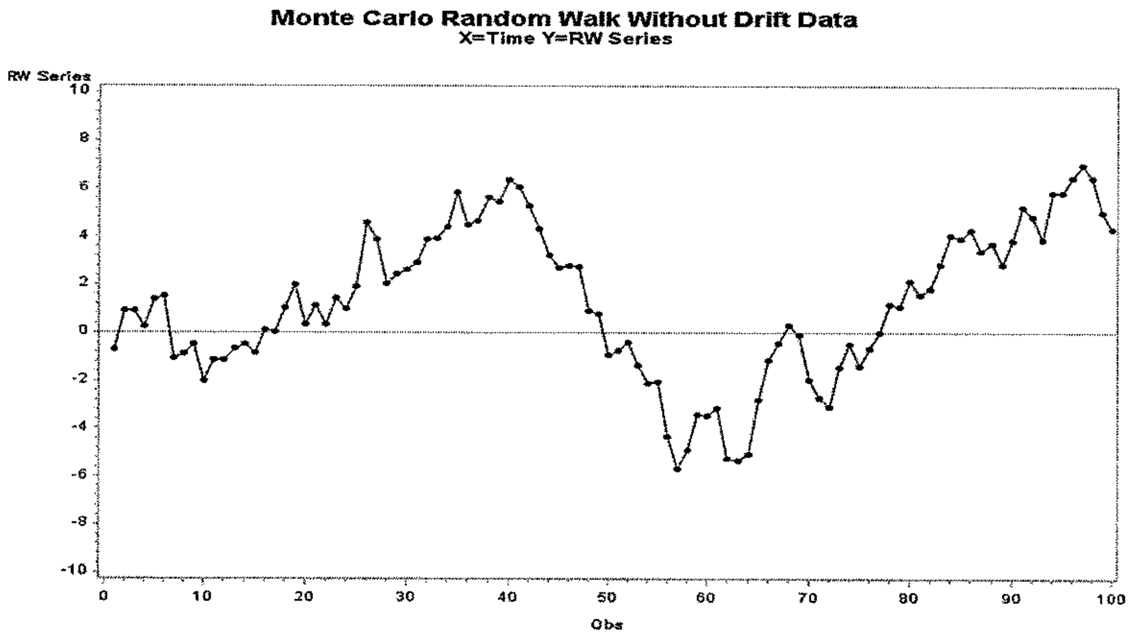


Figure 1

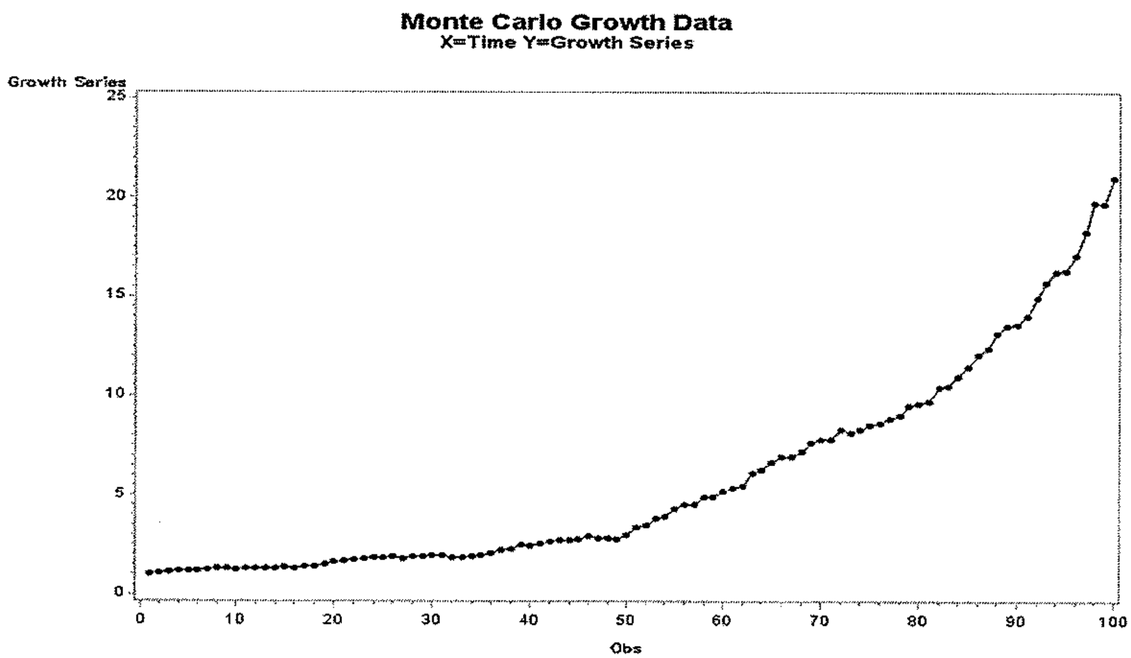


Figure 2

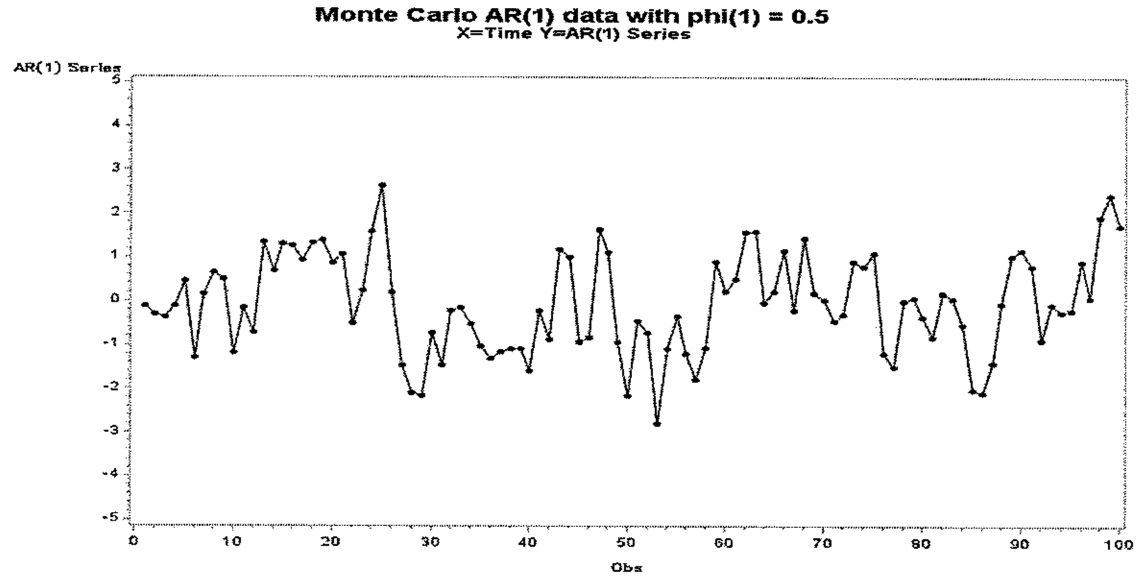


Figure 3

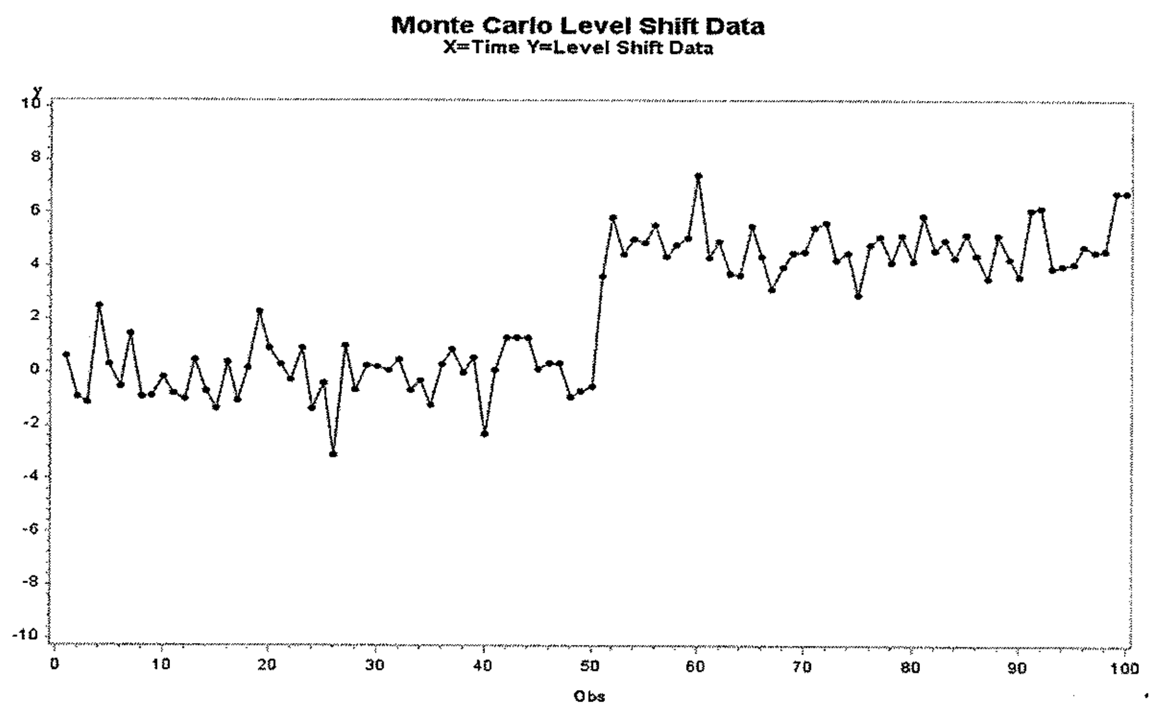


Figure 4

2 18. The treatment that I would use to make the data in Figure 1 stationary is first differencing Δy

2 19. The treatment that I would use to make the data in Figure 2 stationary is Δ (log(y)) : growth rate

2 20. The treatment that I would use to make the data in Figure 3 stationary is no transformation (leave alone)

3 21. The treatment that I would use to make the data in Figure 4 is stationary split the data and model the second half of the data for forecasting purposes.

22. One of the major reasons that we are interested in the Case 3 Augmented Dickey-Fuller tests is because
    a. we want to determine if there is a trend in the data
    b. we want to determine if there is seasonality in the data
    c. we want to know whether the trend in the data is stochastic or deterministic
    d. we want to determine if there is a cycle in the data

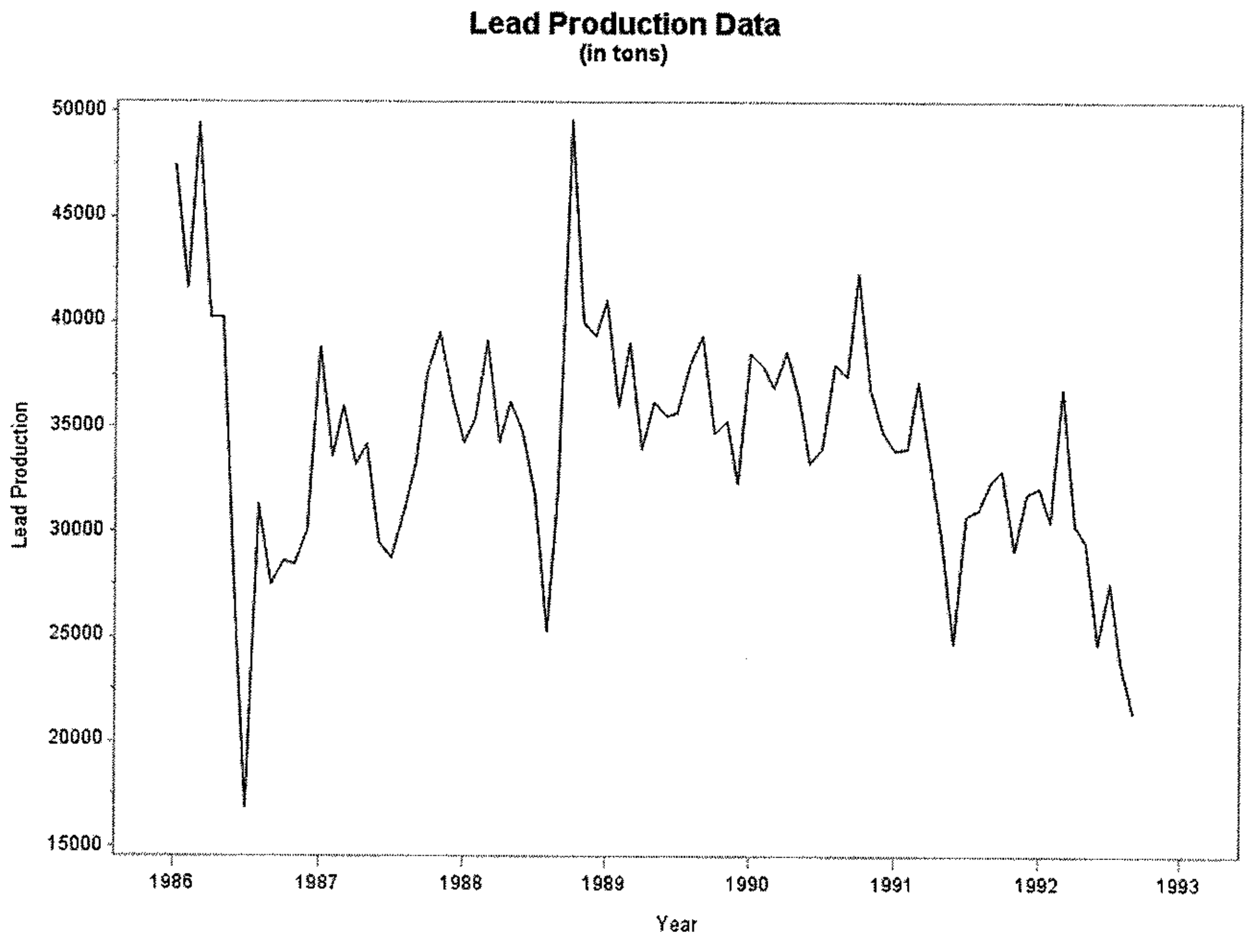
2

23. True / False. The prediction confidence intervals of a stochastic trend grow directly as a function of the square root of the forecast horizon, h.

2

24. Consider the below data in the attached figure. It looks like a
    a. Random Walk without drift
    b. Random Walk with drift
    c. Deterministic Trend
    d. Stationary Time Series

2



Consider the following SAS output that was generated using the "stationary" option in PROC Arima when applied to the above data. Using this output answer the following questions.

(2)

25. Which of the following test equations would you apply to the data in the above graph? Equation __(3)__. Test equation has intercept and no trend.

(1) $\Delta z_t = \theta z_{t-1} + \alpha_1 \Delta z_{t-1} + \alpha_2 \Delta z_{t-2} + \cdots + \alpha_p \Delta z_{t-p} + a_t$

(2) $\Delta z_t = \alpha_0 + \theta z_{t-1} + \gamma t + \alpha_1 \Delta z_{t-1} + \alpha_2 \Delta z_{t-2} + \cdots + \alpha_p \Delta z_{t-p} + a_t$.

(3) $\Delta z_t = \alpha_0 + \theta z_{t-1} + \alpha_1 \Delta z_{t-1} + \alpha_2 \Delta z_{t-2} + \cdots + \alpha_p \Delta z_{t-p} + a_t$.

(3)

26. Given the below SAS output describe to me the proper treatment of the data in the above graph before proceeding to build a Box-Jenkins model for it. Explain your answer in detail. In particular state the **null and alternative hypotheses** of the test you are conducting and your conclusion.

H_0: Data has unit root and should be differenced

H_1: Data is stationary and model data as is.

The probabilities of the Tau statistic at all lags for the single mean case are less than 0.05. Therefore, we reject H_0 and accept H_1 and choose to model the data as is (i.e. with no transformation)

Augmented Dickey-Fuller Unit Root Tests

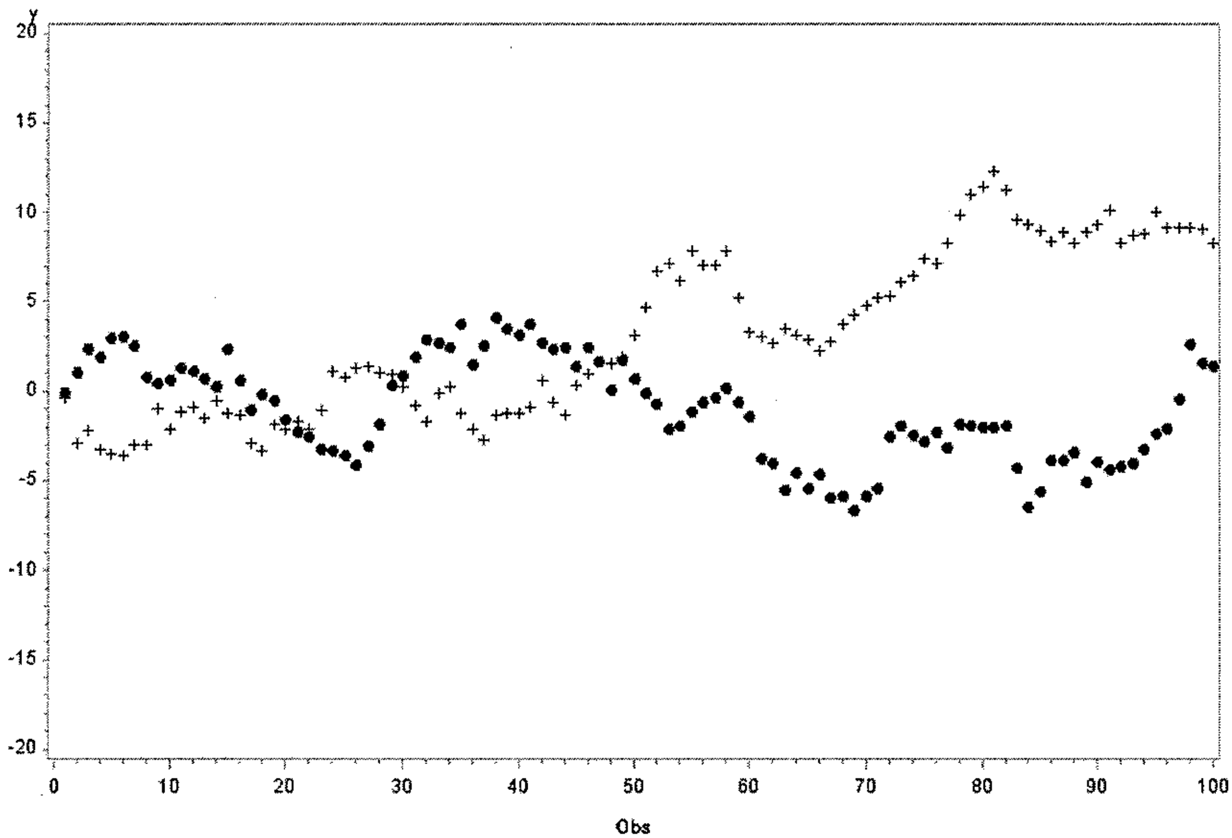
Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-1.5260	0.3893	-1.22	0.2025		
	1	-1.0302	0.4652	-1.06	0.2586		
	2	-1.1196	0.4503	-1.29	0.1816		
	3	-0.8437	0.4982	-1.07	0.2561		
	4	-0.7167	0.5225	-1.11	0.2406		
Single Mean	0	-32.6242	0.0008	-4.48	0.0005	10.27	0.0010
	1	-27.5333	0.0009	-3.50	0.0104	6.35	0.0085
	2	-37.9443	0.0008	-3.96	0.0026	8.25	0.0010
	3	-41.6122	0.0007	-3.43	0.0128	6.18	0.0121
	4	-36.6926	0.0007	-3.00	0.0392	4.90	0.0445
Trend	0	-34.1684	0.0013	-4.58	0.0021	10.51	0.0010
	1	-29.2712	0.0051	-3.61	0.0354	6.53	0.0490
	2	-39.5770	0.0003	-3.98	0.0132	7.94	0.0157
	3	-44.0386	0.0002	-3.48	0.0488	6.07	0.0705
	4	-38.2311	0.0004	-3.03	0.1322	4.59	0.2713

27. Consider the following graph of two time series x and y and the following two estimation equations.

(1) $y_t = \alpha + \beta x_t + a_t$

(2) $\Delta y_t = \alpha + \beta \Delta x_t + a_t$

Two Independent Random Walks without Drift
X=Time Y=Two Independent RW Times Series



Briefly describe to me how you would go about trying to determine if there was a relationship between x and y and why. In your answer be sure and include some discussion of the **spurious correlation problem**. These two series are slow-turning and therefore they are probably I(1). Therefore, examining the relationship between x and y using the levels equation will likely lead to "spurious regressions." (Levels equation = eq.(1)). Therefore, to avoid spurious regression results we should analyze equation (2). This is the appropriate regression to use.

8

28. If you were talking to a lay person, how would tell him/her why it is important to be able to distinguish between time series data that has a "stochastic" trend in it versus time series data that has a deterministic trend in it. A two or three sentence explanation is fine.

Data that has a stochastic trend should be differenced whereas
deterministic trend data should be modeled by a time trend
the two cases.
model, So the statistical models differ in
the prediction confidence intervals differ between
~~differ~~ Also
the two cases. The stochastic trend confidence intervals are
much wider and thus admit more uncertainty than the deterministic
trend confidence intervals.

3