1 Univariate Time Series Analysis

- End goal: Regression model relating a dependent variable to explanatory variables.

- Before doing this we must understand the property of each variable. Thus, this chapter works with single variable, Y_t for $t = 1, .., T$

- New issues with time series data:

1. One time series variable can influence another with a time lag.

2. If the variables are *nonstationary*, a problem known as *spurious regression* may arise.

- We will explain terms like nonstationary, stationary and spurious regression later.

- In this chapter, we use *autoregressive model* (similar to autocorrelated errors of Chapter 5, but applied to Y_t instead of ε_t)

- Using this model, discuss *unit root*.

- If Y_t has a unit root then it is nonstationary.

- *Dickey-Fuller test* is test for unit root

1.1 Trends in Time Series Variables

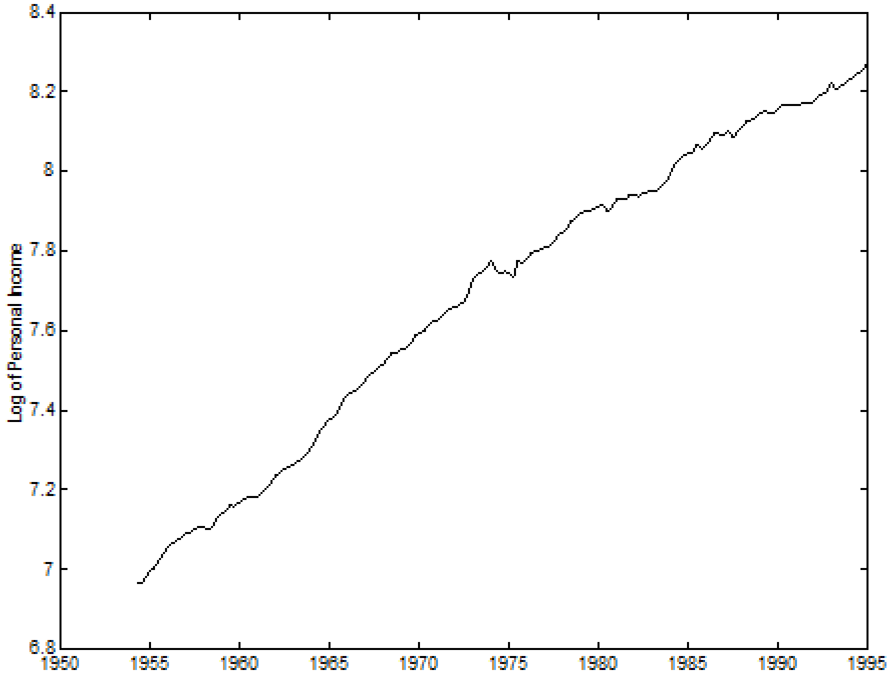
- We have not definited stationary/nonstationary variables, but closely related to the concept of a trend.

- Figure 6.1 plots logarithm of personal income in the U.S. from the first quarter of 1954 through to the last quarter of 1994.

- Note that personal income seems to be increasing over time at a roughly constant rate (although there are fluctuations).

- Many macroeconomic and financial variables (e.g. GDP, the price level, industrial production, consumption, government spending, stock market indices, etc.) exhibit trends of this sort.

Figure 6.1: Time Series Plot of Log of US Personal Income



Differencing

- $\Delta Y_t = Y_t - Y_{t-1}$ is the first difference of Y_t.

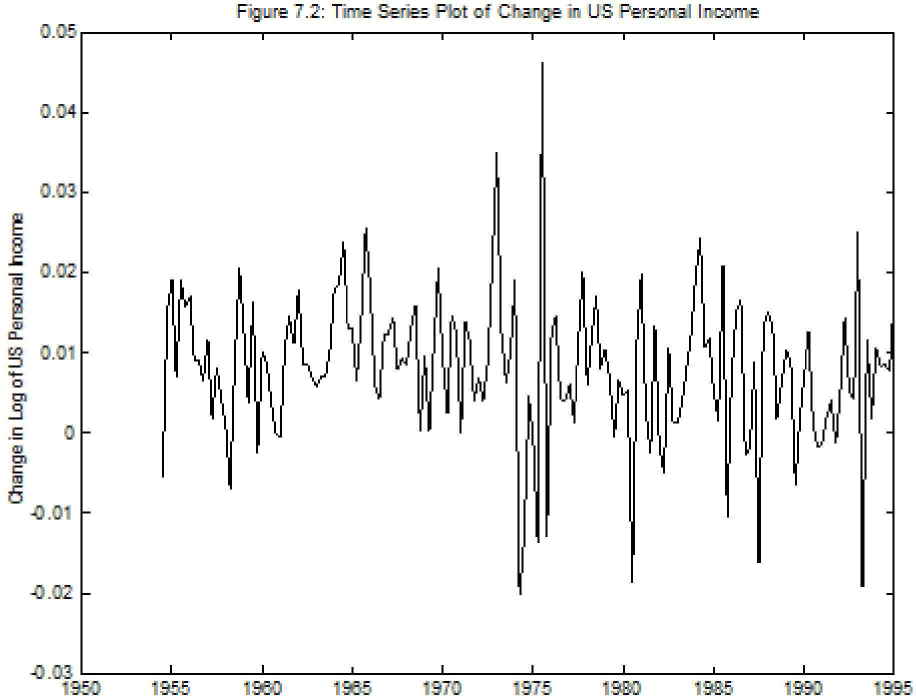
- ΔY_t measures the change or growth in a variable over time.

- If Y_t is variable which has been logged, then $100 \times \Delta Y_t$ measures the percentage change.

- ΔY_t is often called "delta Y" or "the change in Y".

- Figure 7.2 plots ΔY_t for log personal income data

Figure 7.2: Time Series Plot of Change in US Personal Income



- With time series often have correlation across observations.

- Personal income today highly correlated with personal income last quarter (correlation is 0.999716)

- But change in personal income and change in personal income lagged once nearly uncorrelated (-0.00235)

- Many macroeconomic and financial time series exhibit this behavior.

- Y tends to exhibit trend behavior and to be highly correlated over time

- ΔY exhibits no trend behavior and is not highly correlated over time.

- Relate to issue of nonstationarity.

1.2 The Autoregressive Model

- *Autoregressive* model can be used to formalize ideas about trends, stationarity, etc.

- It is a regression model where the explanatory variables are lags of the dependent variable

- "autoregressive" is shortened to "AR".

- Some aspects of AR model discussed in Chapter 6 (errors in a regression had an AR structure – see chapter 6 for some relevant derivations)

- Here AR structure used for Y (not ε).

1.2.1 The AR(1) Model

$$Y_t = \alpha + \rho Y_{t-1} + \varepsilon_t$$

for $t = 2, .., T$.

- Jargon: when Y_t has an AR model, say it is an *autoregressive process*.

- Properties of Y depend on ρ

- Artificially simulate three different time series using $\rho = 0$,$\rho = 0.8$ and $\rho = 1$.

- All three series have the same values for α (i.e. α=0.01) and the same errors.

Figure 7.3: AR(1) Time Series with ρ = 0

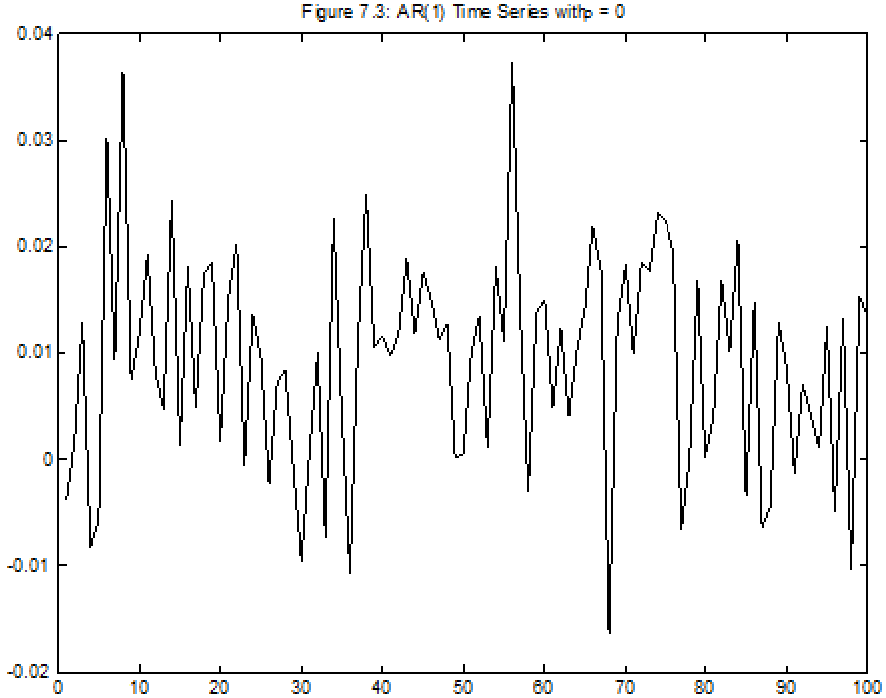


Figure 7.4: AR(1) Time Series with ρ = 0.8

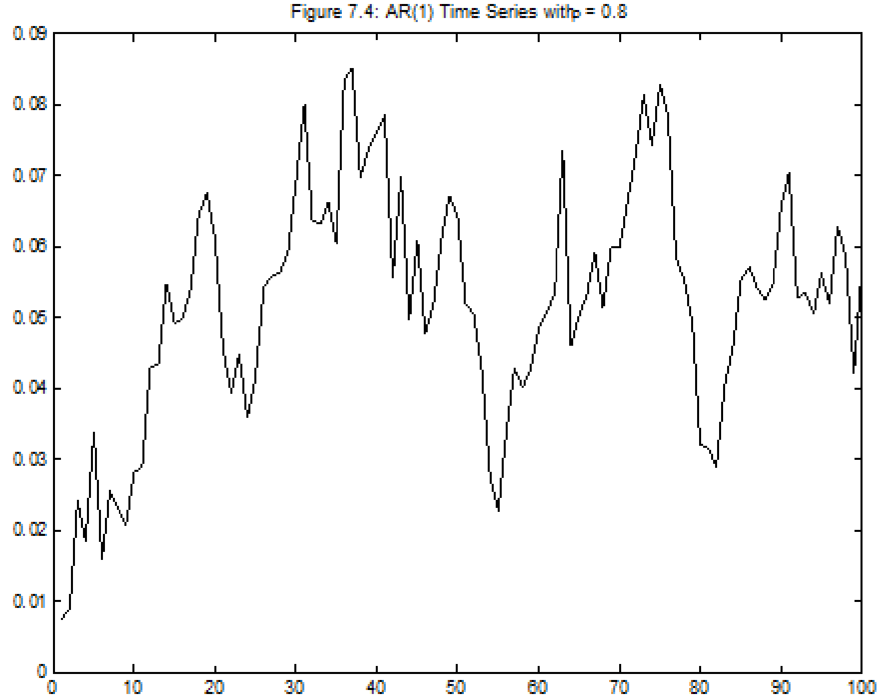
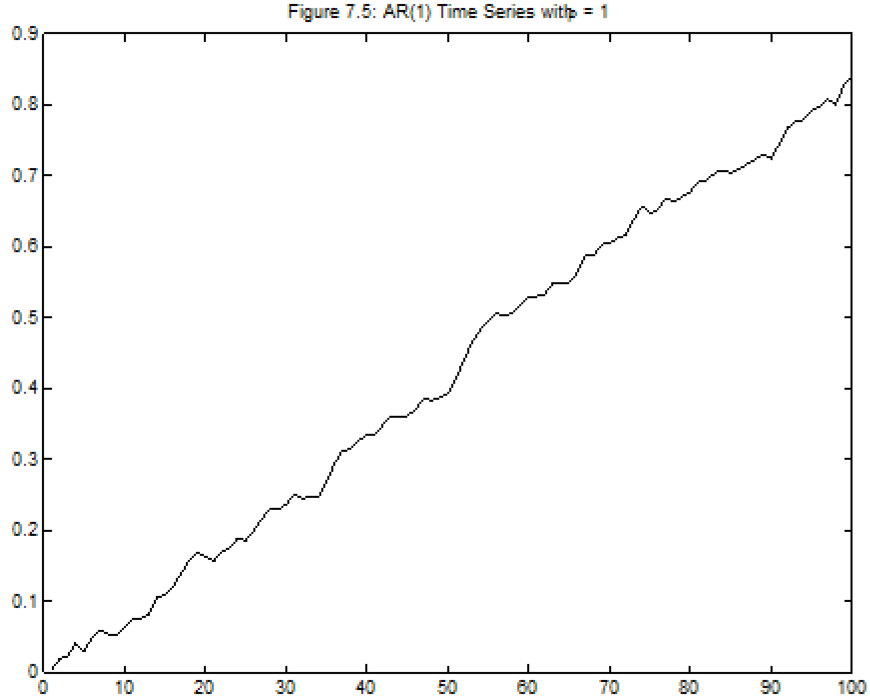


Figure 7.5: AR(1) Time Series with $\rho = 1$



- Figure 7.3 ($\rho = 0$) exhibits random-type fluctuations (like change in personal income)

- Figure 7.5 ($\rho = 1$) exhibits trend behavior (like personal income)

- Figure 7.4 ($\rho = 0.8$) exhibits behavior that is somewhere in-between.

- Figures illustrate the types of behavior that AR(1) models can capture.

- Definition: For the AR(1) model, Y is stationary if $|\rho| < 1$

- Y is nonstationary if $\rho = 1$.

- Note: $|\rho| > 1$, is rarely considered in economics (explosive behaviour).

- To give some math intuition, here is a result from autocorrelated errors part of Chapter 5 (same proof, but uses Y as variable of interest and assumes ε satisfies classical assumptions with $var\left(\varepsilon_t\right) = \sigma^2$):

$$var\left(Y_t\right) = \sigma^2 \sum_{i=0}^{\infty} \rho^{2i}$$

- Note that, if $|\rho| < 1$ we can then write:

$$var\left(Y_t\right) = \frac{\sigma^2}{1 - \rho^2},$$

- But if $\rho = 1$, we cannot (variance is going to infinity, sum does not converge)

- We will not provide proofs/derivations here, but a key aspect of them is: If $\rho = 1$ our basic derivations no longer work (Y will not satisfy classical assumptions, key variances used in our derivations of confidence intervals and hypothesis tests will be going to infinity, etc.).

- Formally, "nonstationary" means "anything that is not stationary".

- Economists focus on the one particular type: unit root nonstationarity.

- Following are ways of thinking about whether, Y, is stationary or has a unit root:

- In the AR(1) model , if $\rho = 1$, then Y has a unit root. If $|\rho| < 1$ then Y is stationary

- If Y has a unit root then its autocorrelations will be near one and will not drop much as lag length increases.

- If Y has a unit root, then it will have a long memory. Stationary time series do not have long memory.

- If Y has a unit root then the series will exhibit trend behavior (especially if α is non-zero).

- If Y has a unit root, then ΔY will be stationary. For this reason, series with unit roots are often referred to as *difference stationary* series.

- A convenient way of writing the AR(1) model (subtract off Y_{t-1} from both sides)

$$Y_t - Y_{t-1} = \alpha + \rho Y_{t-1} - Y_{t-1} + \varepsilon_t$$

which gives

$$\Delta Y_t = \alpha + \phi Y_{t-1} + \varepsilon_t$$

where $\phi = \rho - 1$.

- Note that, if $\rho = 1$, then $\phi = 0$ and ΔY_t fluctuates randomly around α.

- Note we can test $\phi = 0$ to see if a series has a unit root.

- Stationary if $-1 < \rho < 1$ which is equivalent to $-2 < \phi < 0$. This is the *stationarity condition*.

- More jargon: if $\rho = 1$ (or, equivalently, $\phi = 0$) can write:

$$Y_t = \alpha + Y_{t-1} + \varepsilon_t.$$

- This is *random walk* model (more precisely, α is called a drift and, thus, *random walk with drift*.

- Many financial and macroeconomic time series variables look like random walks with drifts.

- The AR(1) model is a regression model. Accordingly, we can use OLS to regress the variable Y on an intercept and lagged Y.

- For personal income, we find $\widehat{\alpha} = 0.039$ and $\widehat{\rho} = 0.996$ (pretty close to random walk with drift)

1.2.2 Extensions of the AR(1) Model

- Autoregressive of order p, AR(p), model:

$$Y_t = \alpha + \rho_1 Y_{t-1} + .. + \rho_p Y_{t-p} + \varepsilon_t$$

for $t = p + 1, .., T$.

- Properties similar to the AR(1) model but are more general.

- When discussing unit root behavior, it is convenient to write model in a different way.

- Subtract Y_{t-1} from both sides of the equation and re-arrange:

$$\Delta Y_t = \alpha + \phi Y_{t-1} + \gamma_1 \Delta Y_{t-1} + .. + \gamma_{p-1} \Delta Y_{t-p+1} + \varepsilon_t,$$

- $\phi, \gamma_1, .., \gamma_{p-1}$ are functions of $\rho_1, .., \rho_p$ from the original AR(p) model.

- For instance, $\phi = \rho_1 + .. + \rho_p - 1$.

- This is identical to the original AR(p) model, but is just written differently.

- $\phi = 0$ implies that the AR(p) time series Y contains a unit root; if $-2 < \phi < 0$, then the series is stationary.

- Note: if $\phi = 0$ then Y_{t-1} will drop out of the equation and only terms involving ΔY or its lags appear in the regression

- if a unit root is present, then the series can be differenced to induce stationarity.

- Adding a *deterministic trend* as an explanatory variable is another way of extending model

- Consider a model where time is explanatory variable:

$$Y_t = \alpha + \delta t + \varepsilon_t,$$

- This regression model yields trend behavior.

- Jargon, δt is *deterministic trend* (unit root series contain *stochastic trend*).

- Can add deterministic trend to AR(1) model:

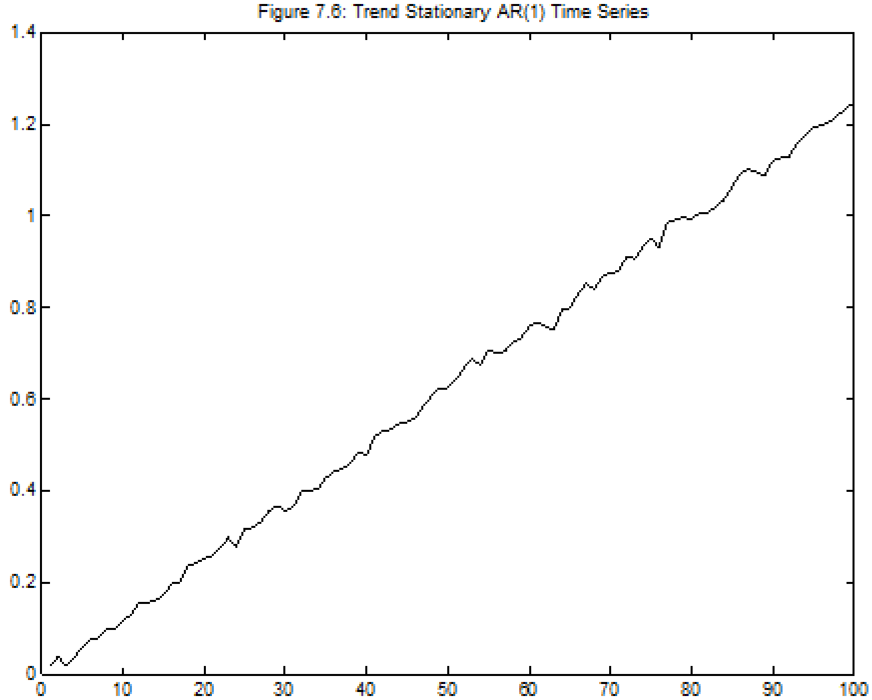
$$Y_t = \alpha + \rho Y_{t-1} + \delta t + \varepsilon_t.$$

- Figure 7.6 which is a time series plot of artificial data generated from the previous model with $\alpha = 0$, $\rho = 0.2$ and $\delta = 0.01$.

- Note it is stationary since $|\rho| < 1$.

- Figure 7.6 looks much like Figure 7.5 (or Figure 7.1).

Figure 7.6: Trend Stationary AR(1) Time Series



- Stationary models with a deterministic trend can yield time series plots that resemble those from unit root.

- Thus, you should not rely on looking at graphs, we need a statistical test for unit root.

- We can add deterministic trend to the AR(p) model to get *AR(p) with deterministic trend model*

$$\Delta Y_t = \alpha + \phi Y_{t-1} + \gamma_1 \Delta Y_{t-1} + .. + \gamma_{p-1} \Delta Y_{t-p+1} + \delta t + \varepsilon_t.$$

1.2.3 AR models: A summary

- The nonstationary time series variables on which we focus are those containing a unit root. These series contain a stochastic trend. But if we difference these time series, the resulting time series will be stationary. For this reason, they are also called *difference stationary*.

- In the stationary time series on which we focus have $-2 < \phi < 0$ in the AR(p) with deterministic trend model. However, these series can exhibit trend behavior through the incorporation of a deterministic trend. In this case, they are referred to as *trend stationary*.

- Now we have to turn to estimation and testing. We will discuss hypothesis testing shortly.

- Estimation of AR models (and extensions): OLS is commonly used (although software packages allow other possible estimation methods)

Example: U.S. Personal Income

Use ΔY_t as dependent variable. Note coefficient on Y_{t-1} is the crucial one for unit root (it is ϕ in previous equations).

Note: if $\phi = 0$ a unit root is present.

Table 7.3: AR(4) with Deterministic Trend Model

Variable	OLS Estimate	t-statistic	P-value
Intercept	0.138	1.279	0.203
Y_{t-1}	-0.018	-1.190	0.236
ΔY_{t-1}	-0.017	-0.217	0.829
ΔY_{t-2}	0.014	0.172	0.863
ΔY_{t-3}	0.130	1.627	0.106
t	0.0001	0.955	0.341

1.3 Testing in the AR(p) with Deterministic Trend Model

- To preview next chapter, with the exception of a case called cointegration, we do not want to include unit root variables in regression models.

- This motivates why we must know if any variable has a unit root. We need a unit root test.

- In previous chapters talked about hypothesis testing to decide whether an explanatory variable should be included, lag length selected, etc.

- Same basic ideas hold.

- However, one important complication occurs in the AR(p) model relating to unit roots.

- To understand it, divide coefficients into two groups: 1) $\alpha, \gamma_1, .., \gamma_{p-1}, \delta$, and 2) ϕ.

- In other words, we consider hypothesis tests involving ϕ independently of those involving the other coefficients.

1.3.1 Testing Involving $\alpha, \gamma_1, .., \gamma_{p-1}$ and δ

- Many methods exist to determine the appropriate lag length in an AR(p) model (textbook discusses the use of *information criteria* , but this course will not cover them).

- But simply looking at t-statistics or F-statistics can be quite informative.

- Such tests involving these coefficients work in the same way as in previous chapters.

- E.g. In Table 7.3 the P-value associated with the coefficients on the lagged ΔY terms are insignificant (so might want to drop these)

- In Table 7.3 the P-value associated with deterministic trend is insignificant (so might want to drop this)

- Alternatively, a common strategy is to choose a maximum lag length, $pmax$, and then sequentially drop lags if the relevant coefficients are insignificant.

A summary of this testing strategy is:

Step 1

Choose the maximum lag length, $pmax$, that seems reasonable.

Step 2

Estimate using OLS the AR($pmax$) with deterministic trend model. If the P-value for testing $\gamma_{p\,\mathsf{max}-1} = 0$ is less than the significance level you choose (e.g. 0.05) then go to Step 5, using $pmax$ as lag length. Otherwise go on to the next step.

Step 3

Estimate the AR($pmax - 1$) model. If the P-value for testing $\gamma_{p\,\mathsf{max}-2} = 0$ is less than the significance level you choose then go to Step 5, using $pmax - 1$ as lag length. Otherwise go on to the next step.

Step 4

Repeatedly estimate lower order AR models until you find an AR(p) model where $\gamma_{p-1} = 0$ is statistically significant (or you run out of lags).

Step 5

Now test for whether the deterministic trend should be omitted; that is, if the P-value for testing $\delta = 0$ is greater than the significance level you choose then drop the deterministic trend variable.

Example: U.S. Personal Income (continued)

We did sequential testing beginning with $pmax = 4$, the model reduces to:

$$\Delta Y_t = \alpha + \phi Y_{t-1} + \varepsilon_t.$$

Results:

Table 7.4: AR(1) Model			
Variable	OLS Estimate	t-statistic	P-value
Intercept	0.039	2.682	0.008
Y_{t-1}	-0.004	-2.130	0.035

1.3.2 Testing Involving ϕ: Unit Root Testing

- Remember if $\phi = 0$, then Y contains a unit root.

- So maybe you can just test $\phi = 0$ in the same manner as you tested the significance of the other coefficients?

- E.g. Table 7.4, the t-statistic for the coefficient ϕ is -2.13. If you get critical values from Student-t statistical tables (or look at P-values), you reject the hypothesis that $\phi = 0$.

- So ϕ is not zero, and, therefore, that Y does not have a unit root??

- THIS IS INCORRECT!

- In hypothesis testing, ϕ is different from other coefficients and we must treat it differently.

- Derivations and proofs a bit hard and will not be provided.

- Textbook provides some hints about why the derivations are different than before.

- Correct test is *Dickey-Fuller test*.

- Dickey-Fuller test uses the familiar t-statistic for testing $\phi = 0$.

- However, critical values for this t-stat not Student-t, but rather *Dickey-Fuller* distribution.

- Another complication: distribution differs between the cases where the AR model does or does not include a deterministic trend.

- Most econometric packages will provide critical values for you (but not Excel)

- Also, I provide some Dickey-Fuller critical values below.

- Note: Some say "Dickey-Fuller test" is for testing for $\phi = 0$ in the AR(1) model and use the term "Augmented Dickey-Fuller test" for testing in the AR(p) model.

- But these are basically the same test so I (and many others) just say "Dickey-Fuller test" for both.

- How to do Dickey Fuller test in practice?

- First estimate the AR(p) model with deterministic trend:

$$\Delta Y_t = \alpha + \phi Y_{t-1} + \gamma_1 \Delta Y_{t-1} + .. + \gamma_{p-1} \Delta Y_{t-p+1} + \delta t + \varepsilon_t$$

and use sequential testing procedures to select a lag length and decide whether the deterministic trend should be included.

- Then, record the t-stat corresponding to the coefficient ϕ and compare to appropriate Dickey-Fuller critical value from Table 7.5.

- Dickey Fuller critical values depend on sample size and whether model has deterministic trend or not.

- Remember $\phi = 0$ implies unit root and $-2 < \phi < 0$, then the series is stationary.

- So if stationary $\widehat{\phi}$ (and, thus, its t-statistic) should be negative.

- Thus, unit root hypothesis is rejected if the t-statistic is more negative than the critical value

Table 7.5: Critical Values for the Dickey-Fuller Test				
	$T = 25$	$T = 50$	$T = 100$	$T = \infty$
AR Model Does Not Have Deterministic Trend				
1% Critical Value	-3.75	-3.59	-3.50	-3.42
5% Critical Value	-2.99	-2.93	-2.90	-2.80
AR Model Does Have Deterministic Trend				
1% Critical Value	-4.38	-4.15	-4.04	-3.96
5% Critical Value	-3.60	-3.50	-3.45	-3.41

- Dickey-Fuller test is the most popular unit root test, however there are many others and many econometrics software packages allow you to do them automatically.

Example: U.S. Personal Income (continued)

- From Table 7.4, Dickey Fuller test statistic is -2.130

- For $T = 163$, Table 7.5 says 5% critical value is between -2.90 and -2.80.

- Test statistic is *not* more negative than the critical value.

- Hence, accept hypothesis that personal income does contain a unit root at the 5% level of significance.

- Note: a formal general definition of stationarity is provided in textbook (but not covered in this course)

- Note: the textbook also have a discussion of volatility/ARCH and GARCH. These topics are not covered in the course (but if you are interested in finance you might want to read them anyway)

1.4 Chapter Summary

1. Regressions with time series variables involve two new issues. First, one variable can influence another with a time lag. Second, if the variables are nonstationary, the spurious regression problem can result. The latter issue will be dealt with in next chapter.

2. Many time series exhibit trend behavior, while their differences do not exhibit such behavior.

3. The autocorrelation function is a common tool for summarizing the relationship between a variable and lags of itself.

4. Autoregressive models are regression models used for working with time series variables. Such models can be written in two ways: one with Y_t as the dependent variable, the other with ΔY_t as the dependent variable.

5. The distinction between stationary and non-stationary models is a crucial one.

6. Unit root nonstationarity is key in economics.

7. If Y_t has unit root then AR(p) model with ΔY_t as the dependent variable can be estimated using OLS. Standard statistical results hold for all coefficients except the coefficient on Y_{t-1}.

8. In the AR(p) model with deterministic trend, sequential hypothesis testing procedures can be used to select lag length and decide whether a deterministic trend should be included.

9. The Dickey-Fuller test is a unit root test. It involves testing whether the coefficient on Y_{t-1} is equal to zero using the t-statistic. The t-statistic does not have a Student-t distribution and critical values must be taken from the Dickey-Fuller statistical tables.