EC220 Review Lectures

Lecture 5

Abhisek Banerjee

07.05.09

Provisional 5 Lecture Topics

1. May 1 (Friday 14:00-16:00):
   Omitted Variable Bias, Including Irrelevant Variables & Heteroscedasticity
2. May 7 (Thursday 11:00-13:00):
   Measurement Error & Simultaneous Equations
3. May 8 (Friday 14:00-16:00):
   Limited Dependent Variable Models
4. May 12 (Tuesday 09:00-11:00):
   Time Series, Dynamic Models & Autocorrelation
5. May 12 (Tuesday 16:00-18:00):
   Nonstationary Time Series

$\rightarrow$ All lectures take place in E171 (New Theater)

Today: Nonstationarity

- Stationary Processes
- Nonstationary Processes
- Spurious Regressions
- Testing for Nonstationarity
- Fitting models with Nonstationary Time Series

Stationary Processes

X_t is stationary if $Cov(X_t, X_{t+s})$ does not depend on t.
Consider:

$$X_t = \beta X_{t-1} + \varepsilon_t$$

where $|\beta| < 1$. Then

$$Corr(X_t, X_{t+s}) = \beta^s$$

This by definition is stationary.

Nonstationary Processes

There are 2 types of nonstationary processes

- Trend Stationary
- Difference Stationary

Trend Stationarity

Trend stationary processes look something like:

$$X_t = \beta_1 + \beta_2 t + \varepsilon_t$$

So we have a deterministic time trend. The term trend stationarity means that the process is stationary around the trend. Such processes can be handled easily by removing the time trend.

$$\begin{aligned} \tilde{X}_t &= X_t - \hat{X}_t \\ \hat{X}_t &= b_1 + b_2 t \end{aligned}$$

Trend Stationarity

Trend stationary processes look something like:

$$X_t = \beta_1 + \beta_2 t + \varepsilon_t$$

So we have a deterministic time trend. The term trend stationarity means that the process is stationary around the trend. Such processes can be handled easily by removing the time trend.

$$\begin{aligned} \tilde{X}_t &= X_t - \hat{X}_t \\ \hat{X}_t &= b_1 + b_2 t \end{aligned}$$

Difference Stationarity

In difference stationary processes, we have that the source if nonstationarity is stochastic rather than deterministic.

$$X_t = X_{t-1} + \varepsilon_t$$

This is a random walk. $E(X_t) = X_0$ and $Var(X_t) = t\sigma^2$. So this is nonstationary.

Adding a constant makes it a random walk with drift.

$$X_t = \beta + X_{t-1} + \varepsilon_t$$

$E(X_t) = t\beta + X_0$ and $Var(X_t) = t\sigma^2$.

Difference Stationarity

In difference stationary processes, we have that the source if nonstationarity is stochastic rather than deterministic.

$$X_t = X_{t-1} + \varepsilon_t$$

This is a random walk. $E(X_t) = X_0$ and $Var(X_t) = t\sigma^2$. So this is nonstationary.
Adding a constant makes it a random walk with drift.

$$X_t = \beta + X_{t-1} + \varepsilon_t$$

$E(X_t) = t\beta + X_0$ and $Var(X_t) = t\sigma^2$.

Difference Stationarity

Suppose we have the following model:

$$X_t = X_{t-1} + \varepsilon_t$$

The first difference of X is:

$$\Delta X_t = X_t - X_{t-1} = \varepsilon_t$$

This is stationary. So differencing removes the nonstationarity. Hence such processes are called difference stationary.

Difference Stationarity

Suppose we have the following model:

$$X_t = X_{t-1} + \varepsilon_t$$

The first difference of X is:

$$\Delta X_t = X_t - X_{t-1} = \varepsilon_t$$

This is stationary. So differencing removes the nonstationarity. Hence such processes are called difference stationary.

Difference Stationarity

Suppose we have the following model:

$$X_t = X_{t-1} + \varepsilon_t$$

The first difference of X is:

$$\Delta X_t = X_t - X_{t-1} = \varepsilon_t$$

This is stationary. So differencing removes the nonstationarity. Hence such processes are called difference stationary.

Difference Stationarity

If X_t is nonstationary and ΔX_t is stationary, then we have a process that is $I(1)$.

But if X_t and ΔX_t are nonstationary and $\Delta^2 X_t$ is stationary,then we have a process that is $I(2)$ and so on.

We call these numbers the orders of integration.

Difference Stationarity

If X_t is nonstationary and ΔX_t is stationary, then we have a process that is $I(1)$.
But if X_t and ΔX_t are nonstationary and $\Delta^2 X_t$ is stationary,then we have a process that is $I(2)$ and so on.
We call these numbers the orders of integration.

Difference Stationarity

If X_t is nonstationary and ΔX_t is stationary, then we have a process that is $I(1)$.
But if X_t and ΔX_t are nonstationary and $\Delta^2 X_t$ is stationary,then we have a process that is $I(2)$ and so on.
We call these numbers the orders of integration.

Spurious Regressions

Why do we care about nonstationarity? Because it can mess up OLS! Refer to the famous simulations by Granger and Newbold.

$$Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$$

where X and Y are independent random walk processes. They found β to be highly significant even though it should be close to 0.

Basically, one gets spurious results. The standard errors and test statistics of the normal t-tests are invalid.

Spurious Regressions

Why do we care about nonstationarity? Because it can mess up OLS! Refer to the famous simulations by Granger and Newbold.

$$Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$$

where X and Y are independent random walk processes. They found β to be highly significant even though it should be close to 0.

Basically, one gets spurious results. The standard errors and test statistics of the normal t-tests are invalid.

Spurious Regressions

Why do we care about nonstationarity? Because it can mess up OLS! Refer to the famous simulations by Granger and Newbold.

$$Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$$

where X and Y are independent random walk processes. They found β to be highly significant even though it should be close to 0.

Basically, one gets spurious results. The standard errors and test statistics of the normal t-tests are invalid.

Spurious Regressions

Why do we care about nonstationarity? Because it can mess up OLS! Refer to the famous simulations by Granger and Newbold.

$$Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$$

where X and Y are independent random walk processes. They found β to be highly significant even though it should be close to 0.
Basically, one gets spurious results. The standard errors and test statistics of the normal t-tests are invalid.

Testing for Nonstationarity

There are a few ways of detecting nonstationarity of which 2 were covered in the problem sets.

1. look at the the time series for X_t
2. look for Durbin Watson statistics close to 0
3. look at autocorrelation function of X_t
4. look at the Augmented Dickey Fuller test statistic

Testing for Nonstationarity

There are a few ways of detecting nonstationarity of which 2 were covered in the problem sets.

1. look at the the time series for X_t
2. look for Durbin Watson statistics close to 0
3. look at autocorrelation function of X_t
4. look at the Augmented Dickey Fuller test statistic

Testing for Nonstationarity

There are a few ways of detecting nonstationarity of which 2 were covered in the problem sets.

1. look at the the time series for X_t
2. look for Durbin Watson statistics close to 0
3. look at autocorrelation function of X_t
4. look at the Augmented Dickey Fuller test statistic

Testing for Nonstationarity

There are a few ways of detecting nonstationarity of which 2 were covered in the problem sets.

1. look at the the time series for X_t
2. look for Durbin Watson statistics close to 0
3. look at autocorrelation function of X_t
4. look at the Augmented Dickey Fuller test statistic

Testing for Nonstationarity

There are a few ways of detecting nonstationarity of which 2 were covered in the problem sets.

1. look at the the time series for X_t
2. look for Durbin Watson statistics close to 0
3. look at autocorrelation function of X_t
4. look at the Augmented Dickey Fuller test statistic

Autocorrelation function

The autocorrelation function is a useful graphical tool for detecting nonstationarity. Suppose we have the following model:

$$X_t = \beta X_{t-1} + \varepsilon_t$$

Then

$$corr(X_t, X_{t+s}) = \beta^s$$

So we can plot the sample autocorrelation function given by

$$\hat{\rho}(s) = \frac{\sum(X_t - \bar{X})(X_{t+s} - \bar{X})}{\sqrt{\sum(X_t - \bar{X})^2 \sum(X_{t+s} - \bar{X})^2}}$$

If the $|\beta| < 1$ then this quantity approximates the true autocorrelation function.

Autocorrelation function

The autocorrelation function is a useful graphical tool for detecting nonstationarity. Suppose we have the following model:

$$X_t = \beta X_{t-1} + \varepsilon_t$$

Then

$$corr(X_t, X_{t+s}) = \beta^s$$

So we can plot the sample autocorrelation function given by

$$\hat{\rho}(s) = \frac{\sum(X_t - \bar{X})(X_{t+s} - \bar{X})}{\sqrt{\sum(X_t - \bar{X})^2 \sum(X_{t+s} - \bar{X})^2}}$$

If the $|\beta| < 1$ then this quantity approximates the true autocorrelation function.

Autocorrelation function

The autocorrelation function is a useful graphical tool for detecting nonstationarity. Suppose we have the following model:

$$X_t = \beta X_{t-1} + \varepsilon_t$$

Then

$$corr(X_t, X_{t+s}) = \beta^s$$

So we can plot the sample autocorrelation function given by

$$\hat{\rho}(s) = \frac{\sum(X_t - \bar{X})(X_{t+s} - \bar{X})}{\sqrt{\sum(X_t - \bar{X})^2 \sum(X_{t+s} - \bar{X})^2}}$$

If the $|\beta| < 1$ then this quantity approximates the true autocorrelation function.

Autocorrelation function

The autocorrelation function is a useful graphical tool for detecting nonstationarity. Suppose we have the following model:

$$X_t = \beta X_{t-1} + \varepsilon_t$$

Then

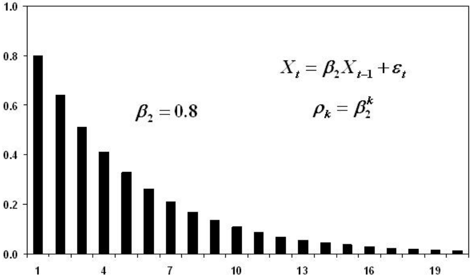
$$corr(X_t, X_{t+s}) = \beta^s$$

So we can plot the sample autocorrelation function given by

$$\hat{\rho}(s) = \frac{\sum(X_t - \bar{X})(X_{t+s} - \bar{X})}{\sqrt{\sum(X_t - \bar{X})^2 \sum(X_{t+s} - \bar{X})^2}}$$

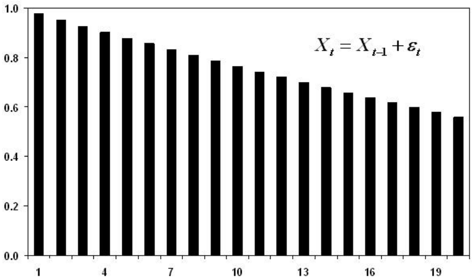
If the $|\beta| < 1$ then this quantity approximates the true autocorrelation function.

Autocorrelation function



Correlogram of an AR(1) process

Autocorrelation function



Correlogram of a random walk

Autocorrelation function

The problems are

- It is hard to differentiate between processes where $\beta_2 = 1$ which is nonstationary and $\beta_2 = 0.95$ which is stationary.
- It is hard to interpret strange graphs which may result from random walk

ADF

Suppose we have the model:

$$X_t = \beta_2 X_{t-1} + \varepsilon_t$$

We would like to test the hypothesis:

$$H_0 : \beta_2 = 1 \text{ vs } H_1 : \beta_2 < 1$$

Note that we are not interested in negative unit roots or $\beta > 1$.
To test this hypothesis, we use a trick - we can transform the equation:

$$\Delta X_t = (\beta_2 - 1)X_{t-1} + \varepsilon_t$$

ADF

Suppose we have the model:

$$X_t = \beta_2 X_{t-1} + \varepsilon_t$$

We would like to test the hypothesis:

$$H_0 : \beta_2 = 1 \text{ vs } H_1 : \beta_2 < 1$$

Note that we are not interested in negative unit roots or $\beta > 1$.
To test this hypothesis, we use a trick - we can transform the equation:

$$\Delta X_t = (\beta_2 - 1) X_{t-1} + \varepsilon_t$$

ADF

Suppose we have the model:

$$X_t = \beta_2 X_{t-1} + \varepsilon_t$$

We would like to test the hypothesis:

$$H_0 : \beta_2 = 1 \text{ vs } H_1 : \beta_2 < 1$$

Note that we are not interested in negative unit roots or $\beta > 1$. To test this hypothesis, we use a trick - we can transform the equation:

$$\Delta X_t = (\beta_2 - 1) X_{t-1} + \varepsilon_t$$

ADF

$$\Delta X_t = (\beta_2 - 1)X_{t-1} + \varepsilon_t$$

We run this regression with OLS and we look at the coefficient of $(\beta_2 - 1)$. The hypothesis now is:

$$H_0 : \beta_2 - 1 = 0 \text{ vs } H_1 : \beta_2 - 1 < 0$$

The test statistic is the Augmented Dickey Fuller statistic. The problem with this statistic is that is has low power. It is hard to differentiate between processes where $\beta_2 = 1$ which is nonstationary and $\beta_2 = 0.95$ which is stationary.
One can allow for a time trend and drift term as well.

ADF

$$\Delta X_t = (\beta_2 - 1)X_{t-1} + \varepsilon_t$$

We run this regression with OLS and we look at the coefficient of $(\beta_2 - 1)$. The hypothesis now is:

$$H_0 : \beta_2 - 1 = 0 \text{ vs } H_1 : \beta_2 - 1 < 0$$

The test statistic is the Augmented Dickey Fuller statistic. The problem with this statistic is that is has low power. It is hard to differentiate between processes where $\beta_2 = 1$ which is nonstationary and $\beta_2 = 0.95$ which is stationary.

One can allow for a time trend and drift term as well.

ADF

$$\Delta X_t = (\beta_2 - 1)X_{t-1} + \varepsilon_t$$

We run this regression with OLS and we look at the coefficient of $(\beta_2 - 1)$. The hypothesis now is:

$$H_0 : \beta_2 - 1 = 0 \text{ vs } H_1 : \beta_2 - 1 < 0$$

The test statistic is the Augmented Dickey Fuller statistic. The problem with this statistic is that is has low power. It is hard to differentiate between processes where $\beta_2 = 1$ which is nonstationary and $\beta_2 = 0.95$ which is stationary.
One can allow for a time trend and drift term as well.

Cointegration

Consider the following true model:

$$Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$$

Suppose that Y and X are both I(1). Can we do OLS? Yes! But if we don't know that this is the true model, how do we know whether the coefficients are valid or not? We can do a test of cointegration

Cointegration

Consider the following true model:

$$Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$$

Suppose that Y and X are both I(1). Can we do OLS? Yes! But if we don't know that this is the true model, how do we know whether the coefficients are valid or not? We can do a test of cointegration

Cointegration

Consider the following true model:

$$Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$$

Suppose that Y and X are both I(1). Can we do OLS? Yes! But if we don't know that this is the true model, how do we know whether the coefficients are valid or not? We can do a test of cointegration

Cointegration

Consider the following true model:

$$Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$$

Suppose that Y and X are both I(1). Can we do OLS? Yes! But if we don't know that this is the true model, how do we know whether the coefficients are valid or not? We can do a test of cointegration

Cointegration

$$Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$$

Two variables are cointegrated if they move together. If this is true, then the error term in the regression should be stationary. That is the idea behind the test for cointegration.

- run the regression with OLS
- grab the OLS residuals
- perform an Augmented Dickey Fuller test on them
- stationarity means that the null is rejected

Note that for cointegration to even be possible, the variables X and Y must be of the same order.

Cointegration

$$Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$$

Two variables are cointegrated if they move together. If this is true, then the error term in the regression should be stationary. That is the idea behind the test for cointegration.

- run the regression with OLS
- grab the OLS residuals
- perform an Augmented Dickey Fuller test on them
- stationarity means that the null is rejected

Note that for cointegration to even be possible, the variables X and Y must be of the same order.

Cointegration

$$Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$$

Two variables are cointegrated if they move together. If this is true, then the error term in the regression should be stationary. That is the idea behind the test for cointegration.

- run the regression with OLS
- grab the OLS residuals
- perform an Augmented Dickey Fuller test on them
- stationarity means that the null is rejected

Note that for cointegration to even be possible, the variables X and Y must be of the same order.

Trend Stationary

For trend stationary, detrending does the job:

$$\begin{aligned} \tilde{X}_t &= X_t - \hat{X}_t \\ \hat{X}_t &= b_1 + b_2 t \end{aligned}$$

This does not work if X is difference stationary (e.g. a random walk). For the random walk, b_2 is likely to be insignificant according to simulation results.

Difference Stationary

One solution to solve nonstationarity problems is just to take first differences:

$$\begin{aligned} Y_t &= \beta_1 + \beta_2 X_t + \varepsilon_t \\ \Delta Y_t &= \beta_2 \Delta X_t + \Delta \varepsilon_t \end{aligned}$$

The problem is that this method only reveals short run dynamics

Difference Stationary

One solution to solve nonstationarity problems is just to take first differences:

$$\begin{aligned} Y_t &= \beta_1 + \beta_2 X_t + \varepsilon_t \\ \Delta Y_t &= \beta_2 \Delta X_t + \Delta \varepsilon_t \end{aligned}$$

The problem is that this method only reveals short run dynamics

Difference Stationary

We would like to uncover long run relationships as well. Suppose we have the following ADL(1,1) model:

$$Y_t = \beta_1 + \beta_2 Y_{t-1} + \beta_3 X_t + \beta_4 X_{t-1} + \varepsilon_t$$

in equilibrium, we have:

$$\bar{Y} = \beta_1 + \beta_2 \bar{Y} + \beta_3 \bar{X} + \beta_4 \bar{X}$$

Rearranging we get the following long run relationship:

$$\bar{Y} = \frac{\beta_1}{1-\beta_2} + \frac{\beta_3 + \beta_4}{1-\beta_2} \bar{X}$$

Difference Stationary

We would like to uncover long run relationships as well. Suppose we have the following ADL(1,1) model:

$$Y_t = \beta_1 + \beta_2 Y_{t-1} + \beta_3 X_t + \beta_4 X_{t-1} + \varepsilon_t$$

in equilibrium, we have:

$$\bar{Y} = \beta_1 + \beta_2 \bar{Y} + \beta_3 \bar{X} + \beta_4 \bar{X}$$

Rearranging we get the following long run relationship:

$$\bar{Y} = \frac{\beta_1}{1-\beta_2} + \frac{\beta_3 + \beta_4}{1-\beta_2} \bar{X}$$

Difference Stationary

We would like to uncover long run relationships as well. Suppose we have the following ADL(1,1) model:

$$Y_t = \beta_1 + \beta_2 Y_{t-1} + \beta_3 X_t + \beta_4 X_{t-1} + \varepsilon_t$$

in equilibrium, we have:

$$\bar{Y} = \beta_1 + \beta_2 \bar{Y} + \beta_3 \bar{X} + \beta_4 \bar{X}$$

Rearranging we get the following long run relationship:

$$\bar{Y} = \frac{\beta_1}{1-\beta_2} + \frac{\beta_3 + \beta_4}{1-\beta_2} \bar{X}$$

Difference Stationary

We can try to uncover both long and short run relationships by having estimating an error correction model. We rearrange the ADL(1,1) (see Prof Dougherty's notes) to get

$$\Delta Y_t = (\beta_2 - 1)(Y_{t-1} - \frac{\beta_1}{1-\beta_2} - \frac{\beta_3+\beta_4}{1-\beta_2} X_{t-1}) + \beta_3 \Delta X_t + \varepsilon_t$$

The first term in brackets contains the long run relationship and the latter term shows the error correction term. We cannot estimate this model straightaway. We could instead use the Engle-Granger 2 step procedure.

Difference Stationary

We can try to uncover both long and short run relationships by having estimating an error correction model. We rearrange the ADL(1,1) (see Prof Dougherty's notes) to get

$$\Delta Y_t = (\beta_2 - 1)(Y_{t-1} - \frac{\beta_1}{1-\beta_2} - \frac{\beta_3 + \beta_4}{1-\beta_2} X_{t-1}) + \beta_3 \Delta X_t + \varepsilon_t$$

The first term in brackets contains the long run relationship and the latter term shows the error correction term. We cannot estimate this model straightaway. We could instead use the Engle-Granger 2 step procedure.

Difference Stationary

- run the long run regression (regress Y on X and a constant)
- grab the residuals which should be stationary since X and Y are cointegrating
- regress ΔY_t on these residuals and ΔX_t

2005 Q8

2005 Q8