Time Series Data

- $y_t = \beta_0 + \beta_1 x_{t1} + \ldots + \beta_k x_{tk} + u_t$
- 1. Basic Analysis

Econometrics V Lecture 9

Time Series vs. Cross Sectional

- Time series data has a temporal ordering, unlike cross-section data
- Will need to alter some of our assumptions to take into account that we no longer have a random sample of individuals
- Instead, we have one realization of a stochastic (i.e. random) process

Econometrics V Lecture 9

Examples of Time Series Models

- A static model relates contemporaneous variables: $y_t = \beta_0 + \beta_1 z_t + u_t$
- A finite distributed lag (FDL) model allows one or more variables to affect *y* with a lag:
  $y_t = \alpha_0 + \delta_0 z_t + \delta_1 z_{t-1} + \delta_2 z_{t-2} + u_t$
- More generally, a finite distributed lag model of order *q* will include *q* lags of *z*

Econometrics V Lecture 9
Finite Distributed Lag Models

- We can call δ_0 the impact propensity – it reflects the immediate change in *y*
- For a temporary, 1-period change, y returns to its original level in period *q*+1
- We can call $\delta_0 + \delta_1 + \ldots + \delta_q$ the long-run propensity (LRP) – it reflects the long-run change in *y* after a permanent change

Econometrics V Lecture 9

Assumptions for Unbiasedness

- Still assume a model that is linear in parameters: $y_t = \beta_0 + \beta_1 x_{t1} + . . . + \beta_k x_{tk} + u_t$
- Still need to make a zero conditional mean assumption: $E(u_t|\boldsymbol{X}) = 0$, $t = 1, 2, …, n$
- Note that this implies the error term in any given period is uncorrelated with the explanatory variables in all time periods

Econometrics V Lecture 9

Assumptions (continued)

- This zero conditional mean assumption implies the x's are strictly exogenous
- An alternative assumption, more parallel to the cross-sectional case, is $E(u_t|\mathbf{x}_t) = 0$
- This assumption would imply the x's are contemporaneously exogenous
- Contemporaneous exogeneity will only be sufficient in large samples

Econometrics V Lecture 9

Assumptions (continued)

- Still need to assume that no *x* is constant, and that there is no perfect collinearity
- Note we have skipped the assumption of a random sample
- The key impact of the random sample assumption is that each u_i is independent
- Our strict exogeneity assumption takes care of it in this case

Econometrics V Lecture 9

Unbiasedness of OLS

- Based on these 3 assumptions, when using time-series data, the OLS estimators are unbiased
- Thus, just as was the case with cross-section data, under the appropriate conditions OLS is unbiased
- Omitted variable bias can be analyzed in the same manner as in the cross-section case

Econometrics V Lecture 9

Variances of OLS Estimators

- Just as in the cross-section case, we need to add an assumption of homoskedasticity in order to be able to derive variances
- Now we assume $\mathrm{Var}(u_t|\boldsymbol{X}) = \mathrm{Var}(u_t) = \sigma^2$
- Thus, the error variance is independent of all the *x*'s, and it is constant over time
- We also need the assumption of no serial correlation: $\mathrm{Corr}(u_t, u_s|\ \boldsymbol{X}) = 0$ for $t \neq s$

OLS Variances (continued)

- Under these 5 assumptions, the OLS variances in the time-series case are the same as in the cross-section case. Also,
- The estimator of σ^2 is the same
- OLS remains BLUE
- With the additional assumption of normal errors, inference is the same

Econometrics V Lecture 9

Trending Time Series

- Economic time series often have a trend
- Just because 2 series are trending together, we can’t assume that the relation is causal
- Often, both will be trending because of other unobserved factors
- Even if those factors are unobserved, we can control for them by directly controlling for the trend

Econometrics V Lecture 9
Trends (continued)

- One possibility is a linear trend, which can be modeled as $y_t = \alpha_0 + \alpha_1 t + e_t$, $t = 1, 2, \ldots$
- Another possibility is an exponential trend, which can be modeled as $\log(y_t) = \alpha_0 + \alpha_1 t + e_t$, $t = 1, 2, \ldots$
- Another possibility is a quadratic trend, which can be modeled as $y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + e_t$, $t = 1, 2, \ldots$

Econometrics V Lecture 9

Detrending

- Adding a linear trend term to a regression is the same thing as using “detrended” series in a regression
- Detrending a series involves regressing each variable in the model on t
- The residuals form the detrended series
- Basically, the trend has been partialled out

Econometrics V Lecture 9

Detrending (continued)

- An advantage to actually detrending the data (vs. adding a trend) involves the calculation of goodness of fit
- Time-series regressions tend to have very high R^2, as the trend is well explained
- The R^2 from a regression on detrended data better reflects how well the x_t's explain y_t

Econometrics V Lecture 9

Seasonality

- Often time-series data exhibits some periodicity, referred to seasonality
- Example: Quarterly data on retail sales will tend to jump up in the 4th quarter
- Seasonality can be dealt with by adding a set of seasonal dummies
- As with trends, the series can be seasonally adjusted before running the regression

Econometrics V Lecture 9

Stationary Stochastic Process

- A stochastic process is stationary if for every collection of time indices $1 \leq t_1 < \ldots < t_m$ the joint distribution of $(x_{t1}, \ldots, x_{tm})$ is the same as that of $(x_{t1+h}, \ldots x_{tm+h})$ for $h \geq 1$
- Thus, stationarity implies that the x_t's are identically distributed and that the nature of any correlation between adjacent terms is the same across all periods

Econometrics V Lecture 9

Covariance Stationary Process

- A stochastic process is covariance stationary if $E(x_t)$ is constant, $Var(x_t)$ is constant and for any t, $h \geq 1$, $Cov(x_t, x_{t+h})$ depends only on h and not on t
- Thus, this weaker form of stationarity requires only that the mean and variance are constant across time, and the covariance just depends on the distance across time

Econometrics V Lecture 9

Weakly Dependent Time Series

- A stationary time series is weakly dependent if x_t and x_{t+h} are "almost independent" as h increases
- If for a covariance stationary process $\text{Corr}(x_t, x_{t+h}) \to 0$ as $h \to \infty$, we'll say this covariance stationary process is weakly dependent
- Want to still use law of large numbers

Econometrics V Lecture 9

A MA(1) Process

- A moving average process of order one [MA(1)] can be characterized as one where $x_t = e_t + \alpha_1 e_{t-1}$, $t = 1, 2, \ldots$ with e_t being an iid sequence with mean 0 and variance σ^2_e
- This is a stationary, weakly dependent sequence as variables 1 period apart are correlated, but 2 periods apart they are not

Econometrics V Lecture 9

An AR(1) Process

- An autoregressive process of order one [AR(1)] can be characterized as one where $y_t = \rho y_{t-1} + e_t$, $t = 1, 2, \ldots$ with e_t being an iid sequence with mean 0 and variance σ_e^2
- For this process to be weakly dependent, it must be the case that $|\rho| < 1$
- $\text{Corr}(y_t, y_{t+h}) = \text{Cov}(y_t, y_{t+h})/(\sigma_y \sigma_y) = \rho_1^h$ which becomes small as h increases

Econometrics V Lecture 9

Trends Revisited

- A trending series cannot be stationary, since the mean is changing over time
- A trending series can be weakly dependent
- If a series is weakly dependent and is stationary about its trend, we will call it a trend-stationary process
- As long as a trend is included, all is well

Econometrics V Lecture 9

Assumptions for Consistency

- Linearity and Weak Dependence
- A weaker zero conditional mean assumption: $E(u_t|\mathbf{x}_t) = 0$, for each t
- No Perfect Collinearity
- Thus, for asymptotic unbiasedness (consistency), we can weaken the exogeneity assumptions somewhat relative to those for unbiasedness

Econometrics V Lecture 9

Large-Sample Inference

- Weaker assumption of homoskedasticity: Var $(u_t|\mathbf{x}_t) = \sigma^2$, for each t
- Weaker assumption of no serial correlation: $E(u_t u_s| \mathbf{x}_t, \mathbf{x}_s) = 0$ for $t \neq s$
- With these assumptions, we have asymptotic normality and the usual standard errors, *t* statistics, *F* statistics and *LM* statistics are valid

Econometrics V Lecture 9
Random Walks

- A random walk is an AR(1) model where $\rho_1 = 1$, meaning the series is not weakly dependent
- With a random walk, the expected value of y_t is always y_0 – it doesn't depend on t
- $\text{Var}(y_t) = \sigma_e^2 t$, so it increases with t
- We say a random walk is highly persistent since $E(y_{t+h}|y_t) = y_t$ for all $h \geq 1$

Econometrics V Lecture 9

Random Walks (continued)

- A random walk is a special case of what’s known as a unit root process
- Note that trending and persistence are different things – a series can be trending but weakly dependent, or a series can be highly persistent without any trend
- A random walk with drift is an example of a highly persistent series that is trending

Econometrics V Lecture 9

Transforming Persistent Series

- In order to use a highly persistent series and get meaningful estimates and make correct inferences, we want to transform it into a weakly dependent process
- We refer to a weakly dependent process as being integrated of order zero, [I(0)]
- A random walk is integrated of order one, [I(1)], meaning a first difference will be I(0)

Time Series Data

- $y_t = \beta_0 + \beta_1 x_{t1} + \ldots + \beta_k x_{tk} + u_t$
- 2. Further Issues

Econometrics V Lecture 9 Testing for AR(1) Serial Correlation

- Want to be able to test for whether the errors are serially correlated or not
- Want to test the null that $\rho = 0$ in $u_t = \rho u_{t-1} + e_t$, $t = 2, \ldots, n$, where u_t is the model error term and e_t is iid
- With strictly exogenous regressors, the test is very straightforward – simply regress the residuals on lagged residuals and use a t-test

Econometrics V Lecture 9 Testing for AR(1) Serial Correlation (continued)

- An alternative is the Durbin-Watson (DW) statistic, which is calculated by many packages
- If the DW statistic is around 2, then we can reject serial correlation, while if it is significantly < 2 we cannot reject
- Critical values are difficult to calculate, making the t test easier to work with

Econometrics V Lecture 9 Testing for AR(1) Serial Correlation (continued)

- If the regressors are not strictly exogenous, then neither the t or DW test will work
- Regress the residual (or *y*) on the lagged residual and all of the *x*'s
- The inclusion of the *x*'s allows each x_{tj} to be correlated with u_{t-1}, so don't need assumption of strict exogeneity

Econometrics V Lecture 9

Testing for Higher Order S.C.

- Can test for AR(*q*) serial correlation in the same basic manner as AR(1)
- Just include *q* lags of the residuals in the regression and test for joint significance
- Can use F test or LM test, where the LM version is called a Breusch-Godfrey test and is $(n-q)R^2$ using R^2 from residual regression
- Can also test for seasonal forms

Econometrics V Lecture 9

Correcting for Serial Correlation

- Start with case of strictly exogenous regressors, and maintain all G-M assumptions except no serial correlation
- Assume errors follow AR(1) so $u_t = \rho u_{t-1} + e_t$, $t = 2, \ldots, n$
- $\text{Var}(u_t) = \sigma^2_e/(1-\rho^2)$
- We need to try and transform the equation so we have no serial correlation in the errors

Econometrics V Lecture 9

Correcting for S.C. (continued)

- Consider that since $y_t = \beta_0 + \beta_1 x_t + u_t$, then $y_{t-1} = \beta_0 + \beta_1 x_{t-1} + u_{t-1}$
- If you multiply the second equation by ρ, and subtract if from the first you get
- $y_t - \rho\, y_{t-1} = (1 - \rho)\beta_0 + \beta_1(x_t - \rho\, x_{t-1}) + e_t$, since $e_t = u_t - \rho\, u_{t-1}$
- This quasi-differencing results in a model without serial correlation

Econometrics V Lecture 9
Feasible GLS Estimation

- Problem with this method is that we don't know ρ, so we need to get an estimate first
- Can just use the estimate obtained from regressing residuals on lagged residuals
- Depending on how we deal with the first observation, this is either called Cochrane-Orcutt or Prais-Winsten estimation

Econometrics V Lecture 9
Feasible GLS (continued)

- Often both Cochrane-Orcutt and Prais-Winsten are implemented iteratively
- This basic method can be extended to allow for higher order serial correlation, AR(q)
- Most statistical packages will automatically allow for estimation of AR models without having to do the quasi-differencing by hand

Econometrics V Lecture 9

Serial Correlation-Robust Standard Errors

- What happens if we don’t think the regressors are all strictly exogenous?
- It’s possible to calculate serial correlation-robust standard errors, along the same lines as heteroskedasticity robust standard errors
- Idea is that want to scale the OLS standard errors to take into account serial correlation

Econometrics V Lecture 9

Serial Correlation-Robust Standard Errors (continued)

- Estimate normal OLS to get residuals, root MSE
- Run the auxiliary regression of x_{t1} on $x_{t2}, \ldots, x_{tk}$
- Form $\hat{a}_t$ by multiplying these residuals with $\hat{u}_t$
- Choose g – say 1 to 3 for annual data, then

$$\hat{v} = \sum_{t=1}^{n} \hat{a}_t^2 + 2\sum_{h=1}^{g} [1 - h/(g+1)] \left(\sum_{t=h+1}^{n} \hat{a}_t \hat{a}_{t-h} \right)$$

and $se(\hat{\beta}_1) = [SE/\hat{\sigma}]^2 \sqrt{\hat{v}}$, where SE is the usual OLS standard error of $\hat{\beta}_j$

T Econometrics V Lecture 9
esting for Unit Roots

- Consider an AR(1): $y_t = \alpha + \rho y_{t-1} + e_t$
- Let H_0: $\rho = 1$, (assume there is a unit root)
- Define $\theta = \rho - 1$ and subtract y_{t-1} from both sides to obtain $\Delta y_t = \alpha + \theta y_{t-1} + e_t$
- Unfortunately, a simple t-test is inappropriate, since this is an I(1) process
- A Dickey-Fuller Test uses the t-statistic, but different critical values

Econometrics V Lecture 9
Testing for Unit Roots (cont)

- We can add *p* lags of Δy_t to allow for more dynamics in the process
- Still want to calculate the t-statistic for θ
- Now it's called an augmented Dickey-Fuller test, but still the same critical values
- The lags are intended to clear up any serial correlation, if too few, test won't be right

Econometrics V Lecture 9

Testing for Unit Roots w/ Trends

- If a series is clearly trending, then we need to adjust for that or might mistake a trend stationary series for one with a unit root
- Can just add a trend to the model
- Still looking at the t-statistic for θ, but the critical values for the Dickey-Fuller test change

- In statistics and econometrics, an **augmented Dickey-Fuller test** (ADF) is a test for a unit root in a time series sample. It is an augmented version of the Dickey-Fuller test for a larger and more complicated set of time series models. The augmented Dickey-Fuller (ADF) statistic, used in the test, is a negative number. The more negative it is, the stronger the rejection of the hypothesis that there is a unit root at some level of confidence

- The testing procedure for the ADF test is the same as for the Dickey-Fuller test but it is applied to the model

$$-\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \ldots + \delta_p \Delta y_{t-p} + \varepsilon_t,$$

- where α is a constant, β the coefficient on a time trend and *p* the lag order of the autoregressive process. Imposing the contraints α = 0 and β = 0 corresponds to modelling a random walk and using the constraint β = 0 corresponds to modelling a random walk with a drift.

- By including lags of the order *p* the ADF formulation allows for higher-order autoregressive processes. This means that the lag length *p* has to be determined when applying the test. One possible approach is to test down from high orders and examine the t-values on coefficients. An alternative approach is to examine information criteria such as the Akaike information criterion, Bayesian information criterion or the Hannon Quinn criterion.

Econometrics V Lecture 9

$$DF_{\tau} = \frac{\hat{\gamma}}{SE(\hat{\gamma})}$$

•The unit root test is then carried out under the null hypothesis γ = 0 against the alternative hypothesis of γ < 0. Once a value for the test statistic

is computed it can be compared to the relevant critical value for the Dickey-Fuller Test. If the test statistic is greater (in absolute value) than the critical value, then the null hypothesis of γ = is rejected and no unit root is present.

4. Other Unit Root Tests:
i. Phillips-Perron Test (PP)

- Phillips and Perron have developed a more comprehensive theory of unit root nonstationarity. The tests are similar to ADF tests, but they incorporate an automatic correction to the DF procedure to allow for autocorrelated residuals.

- The tests usually give the same conclusions as the ADF tests, and the calculation of the test statistics is complex.

Econometrics V Lecture 9

Criticism of Dickey-Fuller and Phillips-Perron-type tests

- Main criticism is that the power of the tests is low if the process is stationary but with a root close to the non-stationary boundary.

  e.g. the tests are poor at deciding if

  $$\phi=1 \text{ or } \phi=0.95,$$

  especially with small sample sizes.
- If the true data generating process (DGP) is

  $$y_t = 0.95y_{t-1} + u_t$$

  then the null hypothesis of a unit root should be rejected.
- One way to get around this is to use a stationarity test as well as the unit root tests we have looked at.

Other unit root tests

ii. Stationarity as the null

- Stationarity tests have

H_0: y_t is stationary

versus H_1: y_t is non-stationary

So that by default under the null the data will appear stationary.

- One such stationarity test is the KPSS test (Kwaitowski, Phillips, Schmidt and Shin, 1992).

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt and Y. Shin, (1992), “Testing the Null Hypothesis of Stationary Against the Alternative of a Unit Root,” *Journal of Econometrics*, 54, 159–178.

- Thus we can compare the results of these tests with the ADF/PP procedure to see if we obtain the same conclusion.

Econometrics V Lecture 9

Other unit root tests

ii. Stationarity as the null

- **A Comparison**

ADF / PP	KPSS
H_0: $y_t \sim I(1)$	H_0: $y_t \sim I(0)$
H_1: $y_t \sim I(0)$	H_1: $y_t \sim I(1)$

- **4 possible outcomes**

a. Reject H_0	and	Do not reject H_0
b. Do not reject H_0	and	Reject H_0
c. Reject H_0	and	Reject H_0
d. Do not reject H_0	and	Do not reject H_0

Other unit root tests

ii. Stationarity as the null

Structural model:

$$\Delta Y_t = \beta + \eta_t + \Delta\xi_t$$

$$Var(\Delta Y_t) = \sigma_\eta^2 + 2\sigma_\xi^2$$

$$\gamma(1) = -\sigma_\xi^2$$

$$\rho(1) = -\frac{\sigma_\xi^2}{(\sigma_\eta^2 + 2\sigma_\xi^2)}$$

$$\gamma(k) = 0, k > 1$$

$$\Delta Y_t \sim MA(1)$$

Reduced form:

$$\Delta Y_t = \beta + \varepsilon_t + \theta\varepsilon_{t-1}$$

$$Var(\Delta Y_t) = \sigma_\varepsilon^2(1+\theta^2)$$

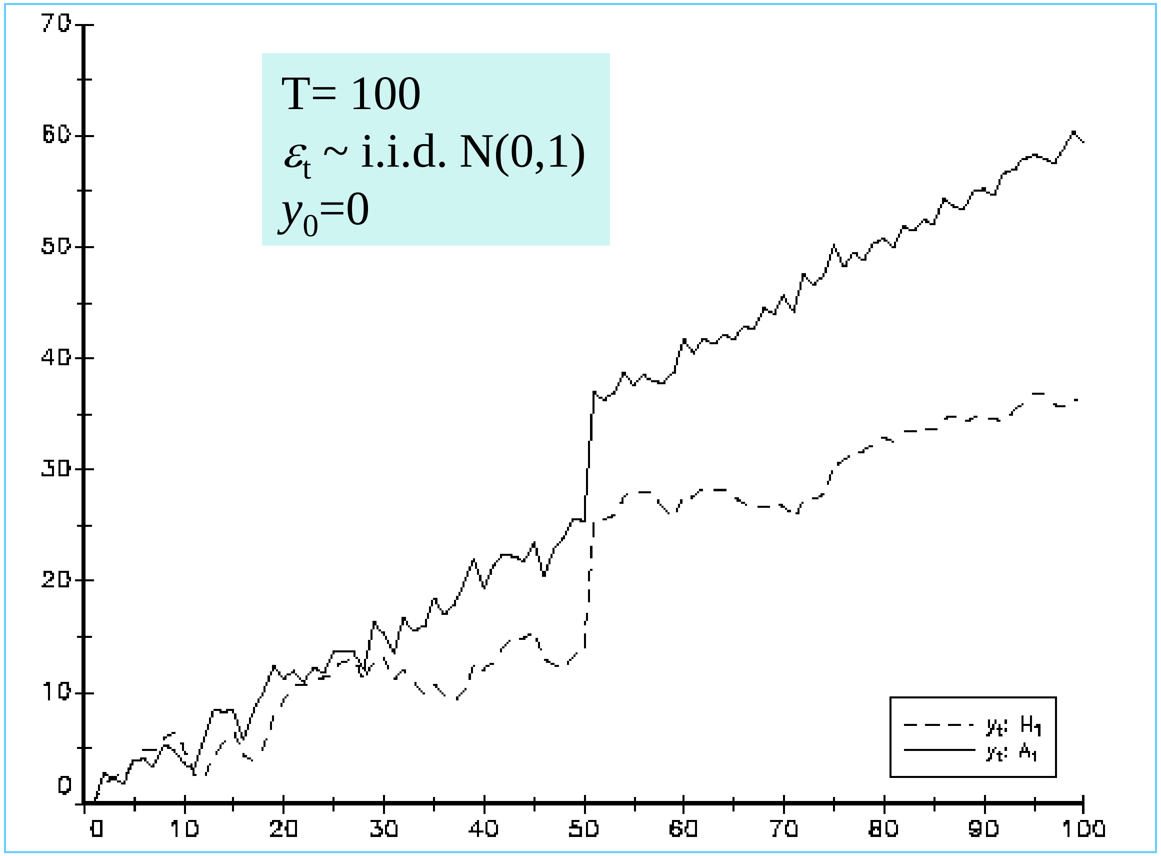
$$\gamma(1) = \theta\sigma_\varepsilon^2$$

$$\rho(1) = \frac{\theta}{(1+\theta^2)}$$

$$\gamma(k) = 0, k > 1$$

Solve θ in terms of the signal-to-noise ratio: $q = \sigma_\eta^2 / \sigma_\xi^2$

Simulated unit root and trend stationary processes with structural break.



H_0: ------
- $a_0 = 0.5$,
- $D_P = 1$ for t = 51 zero otherwise,
- $\mu_1 = 10$.

H_A: ———
- $a_2 = 0.5$,
- $D_L = 1$ for t > 50.
- $\mu_2 = 10$

Power of ADF tests: Rejection frequencies of ADF–tests

	Model: $a_0 = a_2 = 0.5$ and $\mu_2 = 10$		
	1% level	5% level	10% level
ADF–tests	0.004	0.344	0.714
	Model: $a_0 = a_2 = 0.5$ and $\mu_2 = 12$		
ADF–tests	0.000	0.028	0.264

- ADF tests are biased toward nonrejection of the null
- Rejection frequency is inversely related to the magnitude of the shift.
- Perron: estimated values of the autoregressive parameter in the Dickey–Fuller regression was biased toward unity and that this bias increased as the magnitude of the break increased

Econometrics V Lecture 9

Spurious Regression

- Consider running a simple regression of y_t on x_t where y_t and x_t are independent I(1) series
- The usual OLS t-statistic will often be statistically significant, indicating a relationship where there is none
- Called the spurious regression problem

Econometrics V Lecture 9

Cointegration

- Say for two I(1) processes, y_t and x_t, there is a β such that $y_t - \beta x_t$ is an I(0) process
- If so, we say that *y* and *x* are cointegrated, and call β the cointegration parameter
- If we know β, testing for cointegration is straightforward if we define $s_t = y_t - \beta x_t$
- Do Dickey-Fuller test and if we reject a unit root, then they are cointegrated

Econometrics V Lecture 9

Cointegration (continued)

- If β is unknown, then we first have to estimate β, which adds a complication
- After estimating β we run a regression of $\Delta\hat{u}_t$ on $\hat{u}_{t-1}$ and compare t-statistic on $\hat{u}_{t-1}$ with the special critical values
- If there are trends, need to add it to the initial regression that estimates β and use different critical values for t-statistic on $\hat{u}_{t-1}$

Econometrics V Lecture 9
Forecasting

- Once we’ve run a time-series regression we can use it for forecasting into the future
- Can calculate a point forecast and forecast interval in the same way we got a prediction and prediction interval with a cross-section
- Rather than use in-sample criteria like adjusted R^2, often want to use out-of-sample criteria to judge how good the forecast is

Econometrics V Lecture 9

Out-of-Sample Criteria

- Idea is to note use all of the data in estimating the equation, but to save some for evaluating how well the model forecasts
- Let total number of observations be $n + m$ and use n of them for estimating the model
- Use the model to predict the next m observations, and calculate the difference between your prediction and the truth

Econometrics V Lecture 9
Out-of-Sample Criteria (cont)

- Call this difference the forecast error, which is $\hat{e}_{n+h+1}$ for h = 0, 1, …, m
- Calculate the root mean square error and see which model has the smallest, where

$$RMSE = \left(m^{-1} \sum_{h=0}^{m-1} \hat{e}^2_{n+h+1} \right)^{1/2}$$