# Chapter 1 Time Series Basis

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

- 1. Basic Regression Analysis

# 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.

# Examples

- 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*.

# Finite Distributed Lag Models

- We can call $\delta_0$ the impact propensity – it reflects the immediate change in *y*.
- 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.

# 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: $E(u_t|\boldsymbol{X}) = 0$, $t = 1, 2, \ldots,$ n.
- Note that this implies the error term in any given period is uncorrelated with the explanatory variables in all time periods.

# 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 $E(u_t|\mathbf{x}_t) = 0$, which implies the x's are contemporaneously exogenous.
- Contemporaneous exogeneity will only be sufficient in large samples.

# 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.

# 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.

# 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 $\text{Var}(u_t|\boldsymbol{X}) = \text{Var}(u_t) = \sigma^2$ i.e. 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: $\text{Corr}(u_t,u_s|\ \boldsymbol{X})=0$ 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 $\sigma^2$ is the same.
- OLS remains BLUE.
- With the additional assumption of normal errors, inference is the same.

# 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.

# 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 $y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + e_t$, $t = 1, 2, \ldots$

# 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.

# 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 $y_t$.

# 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.

# 2. Large sample properties

# Problems

- Strict exogeneity assumption
- Allow the observations to be correlated across time
- Law of Large Number Theorem and CLT?

# 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.

# 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.

# 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 $\text{Corr}(x_t, x_{t+h}) \rightarrow 0$ as $h \rightarrow \infty$, we'll say this covariance stationary process is weakly dependent.
- We want to still use law of large numbers.

# An MA(1) Process

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

# 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 i.i.d. sequence with mean 0 and variance $\sigma_e^2$.
- 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.

# 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.

# 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.

# 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.

# Random Walks

- A random walk is an AR(1) model:

$$y_t = \rho\, y_{t-1} + e_t$$

where $\rho = 1$. Assume $y_0 = 0$. 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.*
- $\mathrm{Var}(y_t) = \sigma_e^2 t$, so it increases with *t.*
- We say a random walk is highly persistent since $\mathrm{E}(y_{t+h}|y_t) = y_t$ for all $h \geq 1$ since

$$y_{t+h} = e_{t+h} + e_{t+h-1} + \ldots + e_{t+1} + y_t$$

# 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.

# 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).

# 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 i.i.d.
- With strictly exogenous regressors, the test is very straightforward – simply regress the residuals on lagged residuals and use a t-test.

# 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.

# 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.

# Testing for Higher Order S.C.

- Can test for AR($q$) serial correlation in the same 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.

# 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.

# 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 $\rho$, 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.

# Feasible GLS Estimation

- Problem with this method is that we don't know $\rho$, 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.

# 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.

# Serial Correlation-Robust Standard Errors

- What happens if we don't think that 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.

# 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

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

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