Econometrics
Regression Analysis with Time Series Data

João Valle e Azevedo

Faculdade de Economia
Universidade Nova de Lisboa

Spring Semester
Time Series vs. Cross Sectional

\[ y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \ldots + \beta_k x_{tk} + u_t \]

- Time series data is recorded sequentially in time
- We no longer have a random sample of individuals: need new assumptions!
- Instead, we have one realization of a stochastic (i.e. random) process
Time Series

Oil prices ($ per barrel)
Time Series

Figure: Unemployment Rate: U.S. and Europe
Time Series

Figure: Dow Jones Industrial Average
Time Series Example

- **Static Model:** relates contemporaneous variables

\[ y_t = \beta_0 + \beta_1 z_t + u_t \]

- **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

\[ y_t = \alpha_0 + \delta_0 z_t + \delta_1 z_{t-1} + \ldots + \delta_q z_{t-q} + u_t \]

- We can call \( \delta_0 \) the impact propensity: it measures the contemporaneous change in \( y \).
- Given a temporary, 1-period change in \( z \), \( y \) 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 in \( z \).
Assumptions for Time-Series Models

- **Assumption TS.1** (Linearity in parameters) The stochastic process \( \{(x_{t1}, x_{t2}, ..., y_t) : t = 1, 2, ..., n\} \) follows the linear model:

\[
y_t = \beta_0 + \beta_1 x_{t1} + ... + \beta_k x_{tk} + u_t
\]

- \( u_t \) are called the errors or disturbance

- **Assumption TS.2** (No Perfect Collinearity or Absence of Multicollinearity) In the sample, none of the independent variables is a linear combination of others

- **Assumption TS.3** (Zero conditional mean of the error or Strict exogeneity)

\[
E(u_t|X) = 0, t = 1, 2, ..., n
\]

This implies the error term in any given period is uncorrelated with the explanatory variables in all time periods (we do not assume random sampling!)
Unbiasedness of OLS

Under TS.1 through TS.3, OLS estimators are **Unbiased** conditional on $X$, and therefore unconditionally

$$E(u_t|X) = 0, \ t = 1, 2, ..., n$$

is a very strong assumption, often not verified

- Suppose: $CrimeRate_t = \beta_0 + \beta_1 PolicePerCapita_t + u_t$ in a given city
  - $u$ would need to be uncorrelated with current, past and future values of $PolicePerCapita$. We can accept $u$ is uncorrelated with current and past values of the regressor. But clearly, an increase in $u$ today is likely to lead politicians to increase $PolicePerCapita$ in the future! TS.3 fails!

- Suppose: $FarmYield_t = \beta_0 + \beta_1 Labor_t + \beta_2 Rainfall_t + u_t$ for a given farm
  - $u$ would need to be uncorrelated with current, past and future values of $Labor$. But maybe if last year’s $u$ was low (some plague) the farmer will not be able to hire as many workers next year. Ok with $Rainfall$, it will most likely not affect $u$
Unbiasedness of OLS (Cont.)

- We do not worry if $u$ is correlated with past regressors because we can easily solve this problem: just include past regressors, use a distributed lag model.

- But we cannot have $u$ influencing in any way future regressors! (at least to guarantee unbiasedness)

- Omitted variable bias can be analyzed in the same way as for a cross-section.

- An alternative assumption, closer to the cross-sectional case is: $E(u_t|x_t) = 0$. We would say the $x$'s are contemporaneously exogenous. Contemporaneous exogeneity will only be sufficient in large samples.
Variance of OLS Estimators

We need to add an assumption of homoskedasticity in order to be able to derive variances

- **Assumption TS.4** (Homoskedasticity)

  $$Var(u_t|x_t) = \sigma^2 \text{ for } t = 1, 2, ..., n(\text{Compare to TS.4})$$

  - Unlikely to hold in the model:
  
    $$T - billRates_t = \beta_0 + \beta_1 Inflation_t + \beta_2 Deficit_t$$

- **Assumption TS.5** (No Serial Correlation)

  $$Corr(u_t, u_s|X) = Corr(u_t, u_s) = 0 \text{ for } t \neq s$$

The independent variables can be correlated across time!
Variance of OLS Estimators (Cont.)

- With TS.1 through TS.5, the OLS variances in the time-series case are the same as in the cross-section case:

\[
\text{Var}(\hat{\beta}|X) = \sigma^2 (X'X)^{-1} \quad \text{Var}(\hat{\beta}_j) = \frac{\sigma^2}{\text{SST}_j(1 - R_j^2)}
\]

- The estimator of \( \sigma^2 \) is also the same and remains unbiased

- OLS remains BLUE

- With the additional assumption of normal errors, inference is the same

- **Assumption TS.6 (Normality)** The errors are independent of \( X \) and are independent and identically distributed as \( \text{Normal}(0,\sigma^2) \)
Time Series with Trends

- Economic time series often have a trend
- If two series are trending together, we can’t assume that the relation is causal
- We must always control for unobserved factors that can cause the trends. Otherwise we have a **spurious regression problem**
Models for Trends

- One possible control is a **linear trend**

\[ y_t = \alpha_0 + \alpha_1 t + e_t, \ t = 1, 2, ... \]

- Another possibility is an **exponential trend**

\[ \log(y_t) = \alpha_0 + \alpha_1 t + e_t, \ t = 1, 2, ... \]

  where \( \alpha_1 \) is approx. the average growth rate of \( y \)

\[ \log(y_t) - \log(y_{t-1}) \approx (y_t - y_{t-1})/y_{t-1} \]

- Or **quadratic trend**

\[ y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + e_t, \ t = 1, 2, ... \]
Adding trends in a regression

- We should add a trend (usually linear) to the model if either the dependent variable or the independent variables are trending

\[ y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \ldots + \beta_k t + u_t \]

- If Assumptions TS.1 to TS.3 hold in this model, leaving the trend out would in general lead to **biased** estimates of the remaining parameters, specially if the other regressors are trending

- 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 \( t \) and a constant. The **residuals** form the **detrended** series

- Then perform the regression with detrended variables (don’t need intercept, it will equal 0). It will give exactly the same estimates as the regression above
Time Series Analysis

$R^2$ with trending data

- Time-series regressions with trends tend to have a very high $R^2$

- Should therefore look at the $R^2$ from the regression with detrended data

- This $R^2$ better reflects how well the $x_t$’s explain $y_t$

- Can also use an adjusted $R^2$ from the regression with detrended data
## Seasonality

- Often time-series data exhibits seasonal behavior
- Seasonality should be corrected by, e.g., regressing each of the seasonal variables on a set of seasonal dummies
- Can seasonally adjust before running the regression (take the residuals from the previous regression)
- Should look at R-squared only on adjusted data (as for trends)
Important types of Stochastic Processes

- A stochastic process is **stationary** if for every collection of time indices $1 \leq t_1 < \ldots < t_m$ the joint distribution of $(x_{t_1}, \ldots, x_{t_m})$ is the same as that of $(x_{t_1+1}, \ldots, x_{t_m+1})$ for $h \geq 1$

- Otherwise the process is said to be **nonstationary**

- 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 and Weakly Dependent Processes

- 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 $t$.

- 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 $Corr(x_t, x_{t+h}) \to 0$ as $h \to \infty$, we say this covariance stationary process is weakly dependent.

- Need weak dependence to use Laws of Large Numbers and Central Limit Theorems.
An MA(1) Process

A moving average process of order one [MA(1)] satisfies:

\[ y_t = e_t + \alpha_1 e_{t-1}, \ t = 1, 2, \ldots \]

where \( e_t \) is an i.i.d. sequence with mean 0 and variance \( \sigma_e^2 \)

- This is a stationary, weakly dependent sequence as \( y \)'s 1 period apart are correlated, but 2 periods (and further) apart are not
An AR(1) Process

An autoregressive process of order one [AR(1)] satisfies:

\[ y_t = \rho y_{t-1} + e_t, \ t = 1, 2, ... \]

where \( e_t \) is 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 \( |\rho| < 1 \)

\[ \text{Corr}(y_t, y_{t+h}) = \frac{\text{Cov}(y_t, y_{t+h})}{\sigma_y \sigma_y} = \rho_1^h \]

- which becomes small as \( h \) increases (check the derivation!)