1

OLS Regression with Time Series Data

Lecture Plan:

- What are the finite sample properties of OLS regressions with time series data? Times Series (TS) regression model assumptions. Finite sample properties include unbiasedness, and efficiency.
- Gauss-Markov theorem and inference
- Dummy variables and functional form
- Trends and seasonality revisited

2

Finite (Small) Sample Properties of OLS under Classical Assumptions

- In time series analysis, we need three assumptions for OLS estimators to be unbiased

Assumption $TS.1$: Linearity in Parameters

The stochastic process, $\{(x_{t1}, x_{t2}, ..., x_{tk}, y_t) : t = 1, 2, ..., n\}$, follows a model which is linear in parameters

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

where $\{u_t : t = 1, 2, ..., n\}$ is the sequence of errors or disturbances. Here, n is the number of observations (or time periods).

- This assumption is essentially the same as $MLR.1$ (the first cross-sectional assumption).

3

Finite (Small) Sample Properties of OLS under Classical Assumptions

- In the notation, x_{tj}, t denotes the time period and
- j is a label to indicate one of the k explanatory variables.
- In an FDL model, each lag can be defined as a separate variable:

$$x_{t1} = z_t, \quad x_{t2} = z_{t-1} \quad and \quad x_{t3} = z_{t-2}$$

- x_t denotes the set of all explanatory variables at time t.
- Let **X** denote the collection of all independent variables for all time periods.
- It is useful to think of **X** as being an array, with n rows and k columns, as a data matrix.
- the t^th row of **X** is x_t, consisting of all independent variables for time period t.

4

Assumption $TS.2$: No Perfect Collinearity

Assumption $TS.2$: No Perfect Collinearity

In the sample (and therefore in the underlying time series process), no independent variable is constant or a perfect linear combination of the others.

- **No perfect collinearity** is the last assumption required for the OLS estimator to be unbiased.
- Assumption $TS.2$ does allow the explanatory variables, x, to be correlated, but it rules out perfect correlation in the sample.
- Assumption $TS.2$ also covers that there should be sample variability in each explanatory variable.

Assumption $TS.3$: Zero Conditional Mean

Assumption $TS.3$: Zero Conditional Mean (Strict Exogeneity)

For each t, the expected value of the error u_t, given the explanatory variables for all time periods, is zero.

$$E(u_t \mid \mathbf{X}) = 0, \quad t = 1, 2, ..., n.$$

- Assumption $TS.3$ implies that the error at time t, u_t, is uncorrelated with each explanatory variable in every time period.
- The functional relationship must be correctly specified.
- If u_t is **independent** of each column in the matrix $\mathbf{X}$ and $E(u_t) = 0$, then Assumption $TS.3$ automatically holds. (but, this is even stronger assumption).

Assumption $TS.3$: Zero Conditional Mean

- We require u_t to be uncorrelated with the explanatory variables, xs' also dated at time t: in conditional mean terms,

$$E(u_t \mid x_{t1}, x_{t2}, ..., x_{tk}) = E(u_t \mid x_t) = 0$$

- u_t and the explanatory variables are contemporaneously uncorrelated.When the above condition holds, we say that the x_{tj} are **contemporaneously exogenous** (or weakly exogenous).

Contemporaneous Exogeneity

$$Corr(x_{tj}, u_t) = 0, \quad \text{for all } j.$$

- Assumption $TS.3$ requires more than contemporaneous exogeneity.
- u_t must be uncorrelated with x_{sj}, even when $s \neq t$.
- u_t at time t must be uncorrelated with each explanatory variable in every time period

Assumption $TS.3$: Zero Conditional Mean

Strict Exogeneity, $TS.3$

$$E(u_t \mid \mathbf{X}) = 0 \ \text{ implies } \ Corr(x_{sj}, u_t) = 0 \ \text{ for } s \neq t$$

- When $TS.3$ holds, we say that the explanatory variables, x's are **strictly exogenous**.
- The contemporaneous exogeneity is sufficient to prove consistency of the OLS estimator. For OLS to be unbiased, we need the strict exogeneity assumption.
- In the cross-sectional case, we did not explicitly state how u_i, is related to the explanatory variables. The reason this was unnecessary is due to the random sampling assumption ($MLR.2$).
- In a time series context, random sampling is not appropriate, so we must explicitly assume the strict exogeneity.

Assumption $TS.3$: Zero Conditional Mean

- Assumption $TS.3$ puts no restriction on correlation in the independent variables or in the u_t across time.
- Assumption $TS.3$ only says that the average value of u_t is unrelated to the independent variables in all time periods.
- Two leading candidates for failure of this assumption are **omitted variables** and **measurement error**.
- But, the strict exogeneity assumption can also fail for other, less obvious reasons.

Assumption $TS.3$: Zero Conditional Mean

- In the simple static regression model

$$y_t = \beta_0 + \beta_1 z_t + u_t$$

- Assumption $TS.3$ requires not only that u_t and z_t are uncorrelated.
- But u_t is also uncorrelated with past and future values of z: $\{z_{t-1}, z_{t-2}, ...\}$ and $\{z_{t+1}, z_{t+2}, ...\}$.
- This has two implications:
  1. z can have **no lagged effect** on y.
  2. If z does have a lagged effect on y, then we should estimate a distributed lag model.
- **Strict exogeneity** excludes the possibility that changes in the error term today can cause future changes in z.
- This effectively rules out feedback from y on future values of z.

Assumption $TS.3$: Zero Conditional Mean

- For example, consider a simple static model to explain a city's murder rate in terms of police officers per capita:

$$mrdrte_t = \beta_0 + \beta_1 polpc_t + u_t$$

- It may be reasonable to assume that u_t is uncorrelated with $polpc_t$ and even with past values of $polpc_t$.
- For the sake of argument, assume this is the case.
- But suppose that the city adjusts the size of its police force based on past values of the murder rate.
- If this is the case, Assumption $TS.3$ is generally violated.
- This means that, say, $polpc_{t+1}$ might be correlated with u_t (since a higher u_t leads to a higher $mrdrte_t$).
- $u_t \to polpc_{t+1}$ and $mrdrte_t \to polpc_{t+1}$.

Assumption $TS.3$: Zero Conditional Mean

- There are similar considerations in distributed lag models. Usually we do not worry that u_t might be correlated with past z because we are controlling for past z, $\{z_{t-1}, z_{t-2}, ...\}$, in the model.
- But feedback from u to future z is always an issue:
- $u_t \to z_{t+1}, z_{t+2}, ...$
- Explanatory variables that are strictly exogenous cannot react to what has happened to y in the past.
- For example, the amount of rainfall, z_t, at time t, is not influenced by the output during the current or past years, $\{Q_t, Q_{t-1}, Q_{t-2}, ...\}$.
- It also means that rainfall in any future year, $\{z_{t+1}, z_{t+2}, ...\}$, is not influenced by the output during the current or past years.

Assumption $TS.3$: Zero Conditional Mean

- But something like the amount of labor input might not be strictly exogenous, as it is chosen by the farmer, and the farmer may adjust the amount of labor based on last year's yield.
- Therefore, the amount of labor is not strictly exogenous.
- In the social sciences, many explanatory variables may very well violate the strict exogeneity assumption.
- Policy variables, such as growth in the money supply, expenditures on welfare, highway speed limits are often influenced by what has happened to the outcome variable in the past.

Assumption $TS.3$: Zero Conditional Mean

- Even though Assumption $TS.3$ can be unrealistic, we begin with it in order to conclude that the OLS estimators are unbiased.
- Most treatments of static and finite distributed lag models assume $TS.3$ by making the stronger assumption that **the explanatory variables are non-random, or fixed in repeated samples.**
- The new form of the assumption always guarantees assumption $TS.3$.
- But, the non-randomness assumption is obviously false for time series observations, while it isolates the necessary assumption about how u_t and the explanatory variables are related in order for OLS to be unbiased.
- Assumption $TS.3$ has the advantage of being more realistic about the random nature of the x_{tj}.
- in order for OLS to be unbiased, it brings the strict conditions how u_t and the explanatory variables are related ($x \leftrightarrow u$).

Unbiasedness of OLS

Theorem 10.1: Unbiasedness of OLS

Under assumptions $TS.1 - TS.3$, the OLS estimators are unbiased conditional on X, and therefore unconditionally as well:

$$\mathrm{E}(\hat{\beta}_j) = \beta_j, \quad j = 0, 1, 2, \ldots, k$$

- The proof of this theorem is essentially the same as that for Theorem 3.1 in $Ch.3$ for the cross-sectional data.
- Here, the assumption of random sampling is dropped by the assumption that **'for each t, u_t has zero mean given the explanatory variables at all time periods.'**
- If this assumption does not hold, OLS cannot be shown to be unbiased.

Assumption $TS.4$: Homoscedasticity and $TS.5$: No Serial Correlation

- By adding two assumptions let us complete the Gauss-Markov assumptions for time series regressions: **Homoscedasticity** and **No serial correlation**

Assumption $TS.4$: Homoscedasticity

Conditional on $\mathbf{X}$, the variance of u_t is the same for all t: $Var(u_t|\mathbf{X}) = \sigma^2, \quad t = 1, 2, , ..., n.$

Assumption $TS.5$: No Serial Correlation

Conditional on $\mathbf{X}$, the errors in two different time periods are uncorrelated: $Corr(u_t, u_s|\mathbf{X}) = 0$ for all $t \neq s$

$TS.5$: No Serial Correlation

Assumption $TS.5$: No Serial Correlation

Conditional on $\mathbf{X}$, the errors in two different time periods are uncorrelated: $Corr(u_t, u_s|\mathbf{X}) = 0$ for all $t \neq s$

- Assumption $TS.5$ implies that for time t and s, u conditional on $\mathbf{X}$ is uncorrelated.
- We can ignore the conditioning on $\mathbf{X}$ when $\mathbf{X}$ is treated as non-random.

$$Corr(u_t, u_s) = 0 \ for \ all \ t \neq s$$

- When this assumption is not valid, we say that the errors suffer from **serial correlation**, or **autocorrelation**, because they are correlated across time.

Assumption $TS.5$: No Serial Correlation

- Consider the case of errors from adjacent time periods.
- Suppose that, when $u_{t-1} > 0$ then, on average, the error in the next time period, u_t, is also positive. Then $Corr(u_t, u_{t-1}) > 0$ and the errors suffer from serial correlation. The ideal case is that the random errors should be distributed independently of each other.
- Why did we not assume that the errors for different cross-sectional observations are uncorrelated?
- The answer comes from the random sampling assumption: under random sampling, u_i and u_h are independent for any two observations i and h. It can also be shown that this is true, conditional on all explanatory variables in the sample. So, serial correlation is generally an issue in time series regressions.
- When random sampling is not reasonable, it is possible that correlation exists, say, across cities within a state, but as long as the errors are uncorrelated across those cities.

OLS Sampling Variances

Theorem 10.2: OLS Sampling Variances

Under the time series Gauss-Markov assumptions $TS.1$ through $TS.5$,

$$\text{Var}(\hat{\beta}_j \mid \mathbf{X}) = \frac{\sigma^2}{SST_j(1 - R_j^2)}, \quad j = 1, 2, \ldots, k$$

where

$$SST_j = \sum_{t=1}^{n}(x_{tj} - \bar{x}_j)^2$$

x_j is the sample variability of j^{th} independent variable, R_j^2 is the R-squared from the regression of x_j on the other independent variables and the intercept term.

It is the exact variance we derived in Chapter 3 under the cross-sectional regression. Multicollinearity among the explanatory variables, applies immediately to the time series case.

Gauss Markov Theorem

Theorem 10.3: Unbiased Estimation of $\hat{\sigma}^2$

Under Assumptions $TS.1 - TS5$ the estimator $\hat{\sigma}^2 = \frac{SSR}{dof}$ is an unbiased estimator of σ^2, where the degrees of freedom is $df = n - (k+1)$

The usual estimator of the error variance is also unbiased under Assumptions $TS.1$ through $TS.5$, and the Gauss-Markov theorem holds.

Theorem 10.4: Gauss-Markov Theorem

Under Assumptions $TS.1$ through $TS.5$, the OLS estimators are the best linear unbiased estimators (BLUE) conditional on X.

Inference Under the Classical Linear Model Assumptions

- OLS has the same desirable finite (small) sample properties under $TS.1$ through $TS.5$ that it has under $MLR.1$ through $MLR.5$.
- In order to use the usual OLS standard errors, t statistics, and F statistics, we need to add a final assumption that is analogous to the normality assumption we used for cross-sectional analysis.

Assumption $TS.6$: Normality

The errors u_t are independent of X and are independently and identically distributed as $N(0, \sigma^2)$. That is $u_t \sim N(0, \sigma^2)$.

- Assumption $TS.6$ implies $TS.3$, $TS.4$ and $TS.5$.
- In other words, if $TS.6$ holds, $TS.3$, $TS.4$ and $TS.5$ automatically hold.
- But it is stronger because of the independence and normality assumptions.

Inference Under the Classical Linear Model Assumptions

Theorem 10.5: Normal Sampling Distributions

Under Assumptions $TS.1$ through $TS.6$, the CLM assumptions for time series, the OLS estimators are normally distributed, conditional on X. Further, under the null hypothesis, each t statistic has a t distribution, and each F statistic has an F distribution. The usual construction of confidence intervals is also valid.

Inference Under the Classical Linear Model Assumptions

- Theorem 10.5 implies that, when Assumptions $TS.1$ through $TS.6$ hold, everything we have learned about estimation and inference for cross-sectional regressions applies directly to time series regressions. Thus, t statistics can be used for testing statistical significance of individual explanatory variables, and F statistics can be used to test for joint significance.
- The classical linear model assumptions for time series data are much more restrictive than those for the cross-sectional data.
- In particular, the strict exogeneity and no serial correlation assumptions can be unrealistic. Nevertheless, the CLM framework is a good starting point for many applications.

Example 10.1: Static Phillips Curve

- To determine whether there is a tradeoff, on average, between unemployment and inflation
- We can test $H_0 : \beta_1 = 0$ against $H_1 : \beta_1 < 0$.

$$inf_t = \beta_0 + \beta_1 unemp_t + u_t$$

- If the classical linear model assumptions hold, we can use the usual OLS t statistic.

Static Phillips Curve (1948-1996) PHILLIPS.RAW

$$\widehat{\mathrm{inf_t}} = \underset{(1.72)}{1.42} + \underset{(0.289)}{0.468}\, unem_t$$
$$n = 49 \quad R^2 = 0.053 \; \bar{R}^2 = 0.033$$

Example 10.1: Static Phillips Curve

- This equation does not suggest a tradeoff between $unem$ and inf. Because $\hat{\beta}_1$ has an unexpected $(+)$ sign.
- The static Phillips curve is probably not the best model for determining whether there is a short-run tradeoff between inflation and unemployment. The expectations augmented Phillips curve should be preferred.
- Furthermore, the CLM assumptions may not hold in this model.

Example 10.2: Effects of Inflation and Deficits on Interest Rates (US)

- $i3$: the three-month T-bill rate.
- inf : the annual inflation rate based on the consumer price index (CPI).
- def : the federal budget deficit as a percentage of GDP.
- Time Span: Annual, 1948-1996.

Effects of Inflation and Deficits on Interest Rates (INTDEF.RAW)

$$\widehat{\mathrm{i3_t}} = \underset{(0.44)}{1.25} + \underset{(0.076)}{0.613}\, inf_t + \underset{(0.118)}{0.700}\, def_t$$

$$n = 49 \quad R^2 = 0.697 \ \bar{R}^2 = 0.683$$

Example 10.2:

- These estimates show that increases in inflation and the relative size of the deficit work together to increase short-term interest rates, both of which are expected from basic economics.
- A ceteris paribus one percentage point increase in the inflation rate increases $i3$ by 0.613 points.
- Both inf and def are very statistically significant (high t-ratios), assuming, of course, that the CLM assumptions hold.

Functional Form and Dummy Variables

- All of the functional forms in cross-sectional regression can be used in time series regressions.
- The most important of these is the natural logarithm: time series regressions with **constant percentage** effects appear often in applied work.

Example 10.3: Employment and Minimum Wage

- $prepop$: the employment rate in Puerto Rico (ratio of those working to total population)
- $mincov$ = [average minimum wage/average overall wage] $\times$ [proportion of workers covered by minimum wage law]
- $mincov$ measures the importance of the minimum wage relative to average wages.
- $usgnp$: real U.S. gross national product (in billions of dollars).
- Time Span : 1950-1987.

Employment and Minimum Wage (data = prminwge)

$$\log(\widehat{\mathrm{prepop}})_t = -\underset{(0.77)}{1.05} - \underset{(0.065)}{0.154}\, log(mincov)_t - \underset{(0.118)}{0.012}\, log(usgnp)_t$$

$$n = 38 \quad R^2 = 0.661 \ \bar{R}^2 = 0.641$$

Example 10.3: Employment and Minimum Wage

- The estimated elasticity of $prepop$ with respect to $mincov$ is 0.154, and it is statistically significant.
- Therefore, a higher minimum wage lowers the employment rate, something that classical economics predicts.
- The $usngp$ variable is not statistically significant, but this changes when we account for a time trend in the next section.

FDL Model and Logarithmic Functional Form

- We can use logarithmic functional forms in finite distributed lag models, FDL, too.
- For example, for quarterly data, suppose that money demand (M) and gross domestic product (GDP) are related by

FDL(4)

$$log(M_t) = \alpha_0 + \delta_0 log(GDP)_t + \delta_1 log(GDP)_{t-1} + \delta_2 log(GDP)_{t-2}$$

$$+ \delta_3 log(GDP)_{t-3} + \delta_4 log(GDP)_{t-4} + u_t$$

FDL Model and Logarithmic Functional Form

- The impact propensity in this equation, δ_0, is also called the **short-run elasticity**: it measures the immediate percentage change in money demand given a 1% increase in GDP.
- The long-run propensity, $\delta_0 + \delta_1 + \delta_2 + \delta_3 + \delta_4$, is sometimes called the **long-run elasticity**: it measures the percentage increase in money demand after four quarters given a permanent 1% increase in GDP.

Time Series Analysis and Dummy (Binary) Variables

- Binary or dummy independent variables are also quite useful in time series applications.
- Since the unit of observation is time, a dummy variable represents whether, in each time period, a certain event has occurred.
- For example, for annual data, we can indicate in each year whether a political party is in power which is unity if an X party is in power, and zero otherwise.
- Often dummy variables are used to isolate certain periods that may be systematically different from other periods covered by a data set such as wars, crisis, earthquakes.
- Binary explanatory variables are the key component in what is called an **event study**. In an event study, the goal is to see whether a particular event influences some outcome.

Example 10.4: Effects Of Personal Tax Exemption on Fertility Rates, (US, 1913-1984)

- gfr: the number of children born to every 1,000 women of childbearing age
- pe: the average real dollar value of the personal tax exemption when having a child.
- $ww2$: takes on the value unity (1) during the years 1941 through 1945, when the United States was involved in World War II, otherwise 0
- $pill$: takes on the value unity from 1963 on, when the birth control pill was made available for contraception.

Tax Exemption and Fertility (FERTIL3.RAW)

$$\widehat{\mathrm{gfr}}_t = \underset{(3.21)}{98.68} + \underset{(0.030)}{0.083}\, pe_t - \underset{(7.46)}{24.24}\, ww2_t - \underset{(4.08)}{31.59}\, pill_t$$

$$n = 72 \quad R^2 = 0.473 \;\; \bar{R}^2 = 0.450$$

Example 10.4: Tax Exemption and Fertility

- Each variable is statistically significant at the 1% level against a two-sided alternative.
- The fertility rate was lower during World War II: given pe, there were about 24 fewer births for every 1,000 women of childbearing age, which is a large reduction. (From 1913 through 1984, gfr ranged from about 65 to 127.)
- Similarly, the fertility rate has been substantially lower since the introduction of the birth control pill.
- The variable of economic interest is pe. The average pe over this time period is 100.4 dollars, ranging from zero to 243.83.

Example 10.4: Tax Exemption and Fertility

- The coefficient on pe implies that a 12-dollar increase in pe increases gfr by about one birth per 1,000 women of childbearing age (12 x 0.083 = 1).
- The fertility rate may react to changes in pe with a lag.
- Estimating a distributed lag model with two lags

Tax Exemption and Fertility: FDL(2)

$$\widehat{\mathrm{gfr}}_t = \underset{(3.28)}{95.87} + \underset{(0.126)}{0.073}\, pe_t - \underset{(0.1557)}{0.0058}\, pe_{t-1}$$
$$+ \underset{(0.126)}{0.034}\, pe_{t-2} - \underset{(10.73)}{22.13}\, ww2_t - \underset{(3.98)}{31.30}\, pill_t$$

$$n = 70 \quad R^2 = 0.499 \;\; \bar{R}^2 = 0.459$$

- n=70 because we lose two when we lag pe twice.

Example 10.4: Tax Exemption and Fertility

- The coefficients on the pe variables are estimated very imprecisely, and each one is individually insignificant.
- It turns out that there is substantial correlation between pe_t, pe_{t-1} and pe_{t-2} and this multicollinearity makes it difficult to estimate the effect at each lag.
- However, pe_t, pe_{t-1} and pe_{t-2} are jointly significant: the F statistic has a p-value of 0.012.
- Thus, pe does have an effect on gfr , but we do not have good enough estimates to determine whether it is contemporaneous or with a one or two-year lag (or some of each).
- Actually, pe_{t-1} and pe_{t-2} are jointly insignificant in this equation (p-value 0.95) So, we would be justified in using the static model.

Example 10.4: Tax Exemption and Fertility

- But for illustrative purposes, let us obtain a confidence interval for the long-run propensity in this model.
- The long run propensity of the FDL(2) model is $0.073 - 0.0058 + 0.034 = 0.101$.
- To obtain the standard error of the estimated LRP, we use the trick suggested in Chapter 4.
- $\theta_0 = \delta_0 + \delta_1 + \delta_2$ is the **LRP**
- Substitute for $\delta_0 = \theta_0 - \delta_1 - \delta_2$ in the model

$$\begin{aligned} gfr_t &= \alpha_0 + \delta_0 pe_t + \delta_1 pe_{t-1} + \delta_2 pe_{t-2} + \ldots \\ &= \alpha_0 + (\theta_0 - \delta_1 - \delta_2) pe_t + \delta_1 pe_{t-1} + \delta_2 pe_{t-2} + \ldots \\ &= \alpha_0 + \theta_0 pe_t + \delta_1 (pe_{t-1} - pe_t) + \delta_2 (pe_{t-2} - pe_t) + \ldots \end{aligned}$$

Example 10.4: Tax Exemption and Fertility

- Running this last regression gives $\hat{\theta}_0$= 0.101 as the coefficient on pe_t and stderror($\hat{\theta}_0$) $= 0.030$. Therefore, the t statistic is about 3.37, so it is statistically significant.
- Even though none of the δ is individually significant, the LRP is very significant. The 95% confidence interval for the θ_0 is about 0.041 to 0.160.

Using Trend in Time Series Analysis

- Many economic time series have a common tendency of growing over time.
- We must recognize that some time series contain a **time trend** in order to draw causal inference using time series data.
- Ignoring the fact that two sequences are trending in the same or opposite directions can lead us to falsely conclude that changes in one variable are actually caused by changes in another variable.
- In many cases, two time series processes appear to be correlated only because they are both trending over time for reasons related to other unobserved factors.
- What kind of statistical models adequately capture trending behavior?

Using Trend in Time Series Analysis

- One popular model is the **linear time trend** model.

Linear Time Trend Model

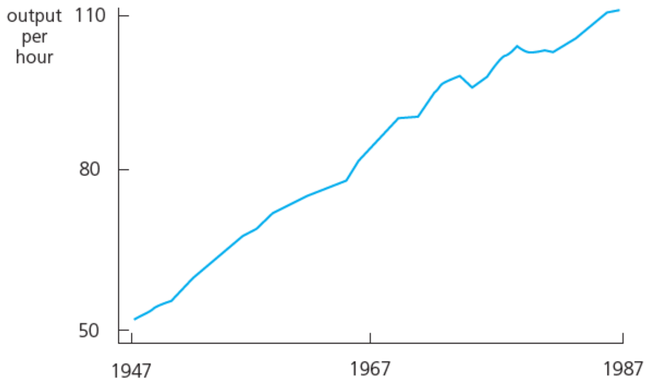
$$y_t = \alpha_0 + \alpha_1 t + e_t, \qquad t = 1, 2, \ldots$$

- where (e_t) is an **independent and identically distributed, iid** sequence.

$$E(e_t) = 0 \qquad Var(e_t) = \sigma_e^2$$

- Note how the parameter α_1 multiplies time, t, resulting in a linear time trend.
- In this model, holding all other factors (those in e_t) fixed, α_1 measures the change in y from one period to the next due to the passage of time.

1947-1987 Output per labor hour in the US: 1977=100



An example of linear trend

Using Trend in Time Series Analysis

- When $\Delta e_t = 0$,
$$\Delta y_t = y_t - y_{t-1}$$
- Another way to think about a sequence that has a linear time trend is that its average value is a linear function of time:
$$E(y_t) = \alpha_0 + \alpha_1 t.$$
- If $\alpha_1 > 0$, then, on average, y_t is growing over time and therefore has an upward trend.
- If $\alpha_1 < 0$, then y_t has a downward trend.
- The values of y_t do not fall exactly on the line due to randomness, but the expected values are on the line.
- For e_t takes on random values, y_t fluctuates around the linear time trend.

Using Trend in Time Series Analysis

- Unlike the mean, the variance of y_t is constant across time:
$$Var(y_t) = Var(y_{t-1}) = \sigma_e^2$$
- If $\{e_t\}$ is an i.i.d. sequence, then $\{y_t\}$ is an independent, though not identically, distributed sequence.
- A more realistic characterization of trending time series allows e_t to be correlated over time, but this does not change the flavor of a linear time trend.

Using Trend in Time Series Analysis

- Many economic time series are better approximated by an **exponential trend**, which follows when a series has the same average growth rate from period to period.
- In practice, an exponential trend in a time series is captured by modeling the natural logarithm of the series as a linear trend (assuming that $y_t > 0$):

Exponential Trend Model

$$y_t = exp(\beta_0 + \beta_1 t + e_t)$$

- β_1 is called the growth rate in y from period $t-1$ to period t. Taking the natural logarithm of both sides:
$$log(y_t) = \beta_0 + \beta_1 t + e_t$$
$$\Delta log(y_t) \approx \frac{y_t - y_{t-1}}{y_{t-1}}$$

Exponential Trend Model

$$y_t = exp(\beta_0 + \beta_1 t + e_t)$$

- Taking the natural logarithm of both sides:

$$log(y_t) = \beta_0 + \beta_1 t + e_t$$

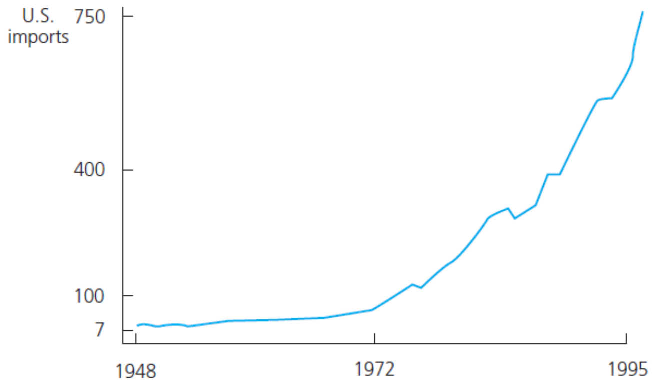
$$\Delta log(y_t) \approx \frac{y_t - y_{t-1}}{y_{t-1}}$$

- Setting $\Delta e_t = 0$,

$$\Delta log(y_t) = \beta_1$$

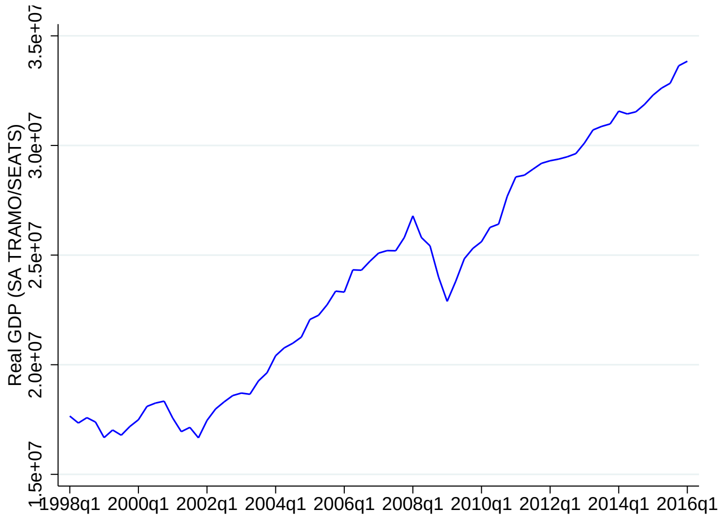
- In other words, β_1 is approximately the average per period growth rate in y_t.
- For example, if t denotes year and $\beta_1 = 0.027$, then y_t grows about 2.7% per year on average.

An example of exponential trend



Nominal U.S. imports (in billions of U.S. dollars)

Exponential Trend Example: Turkish Real GDP



Turkish Real GDP 1998q1-2016q1 (seasonally adjusted)

Exponential Trend: Turkish Real GDP

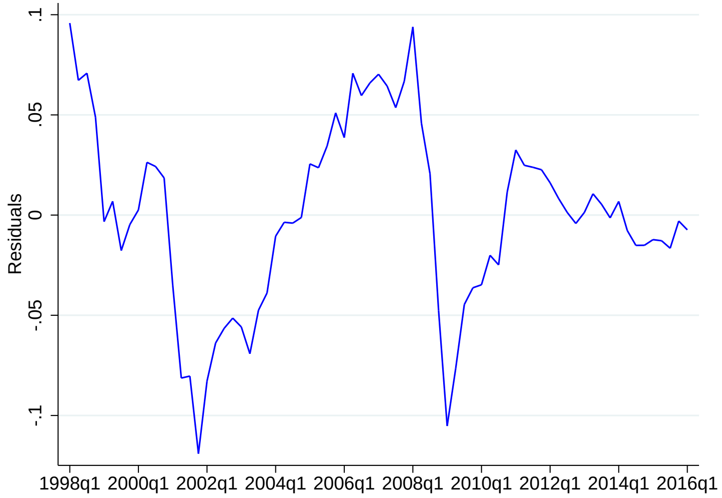
Regression of log(Real GDP) on trend in equation form:

$$log(\widehat{RealGDP}) = 16.58 + 0.01trend$$

Interpretation: Average growth rate at each quarter is approximately 1%. This corresponds to an annual growth rate of 4% on average.

Exponential Trend: Turkish Real GDP

Residuals from the exponential trend regression



Using Trend in Time Series Analysis

- Although linear and exponential trends are the most common, time trends can be more complicated.
- When the slope of the trend changes over time (increasingly or decreasingly), t^2 can be added as a regressor to the model:

Quadratic Trend Model

$$y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + e_t$$

- If α_1 and α_2 are positive, the slope of the trend is increasing with t .
- If $\alpha_1 > 0$ and $\alpha_2 < 0$, the trend has a hump shape. The slope of the trend is decreasing with t.

$$\frac{\Delta y_t}{\Delta t} \approx \alpha_1 + 2\alpha_2 t$$

Using Trending Variables in Regression Analysis

- The trending variables in regression analysis do not violate the assumptions, $TS.1$ through $TS.6$.
- However, we must be careful to allow for the fact that unobserved, trending factors that affect y_t might also be correlated with the explanatory variables. This is called the **spurious regression**.
- Adding a time trend eliminates this problem.

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \beta_3 t + u_t$$

- Allowing for the trend (t) in this equation recognizes that y_t may be growing ($\beta_3 > 0$) or shrinking ($\beta_3 < 0$) over time for reasons essentially unrelated to x_{t1} and x_{t2}.
- Omitting t as a regressor,when necessary, causes the omitted variable bias.

Example 10.7: Housing Investment and Prices

- $invpc$: real per capita housing investment and $price$: a housing price index (1982=1)
- There might be a superiors regression. When we add t, the relationship between the variables disappeared.

Regression Without Trend

$$\widehat{\log(\text{invpc})} = \underset{(0.43)}{-5.50} + \underset{(0.382)}{1.241}\, log(price)$$
$$n = 42 \quad R^2 = 0.208 \ \bar{R}^2 = 0.189$$

Regression With Trend

$$\widehat{\log(\text{invpc})} = \underset{(0.136)}{-0.913} - \underset{(0.679)}{0.381}\, log(price) + \underset{(0.0035)}{0.0098}\, t$$
$$n = 42 \quad R^2 = 0.341 \ \bar{R}^2 = 0.307$$

Using Trending Variables in Regression Analysis

- Including a time trend in a regression model is the same as **detrending** the original data series before using them in regression analysis. For concreteness, consider the following three models:
  1. $y_t = \alpha_0 + \alpha_1 t + u_t$ MODEL A
  2. $x_{t1} = \theta_0 + \theta_1 t + e_t$ MODEL B
  3. $x_{t2} = \gamma_0 + \gamma_1 t + h_t$ MODEL C
- Regress each of y_t, x_{t1} and x_{t2} on a constant and the time trend t and save the residuals of each model (Detrending).
- If we regress the residuals of MODEL A on the residuals of MODEL B and MODEL C, the slope estimates of primary interest in this model (No intercept is necessary, but including an intercept affects nothing: the intercept will be estimated to be zero.)

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \beta_3 t + error$$

- are exactly as the same as the slope estimates $\hat{\beta}_1$ and $\hat{\beta}_2$ coming from a regression with time trend.

Using Trending Variables in Regression Analysis

- R^2 in time series regressions are often very high, compared with typical R^2 for cross-sectional data
- One reason is that time series data often come in aggregate form.
- Aggregates are often easier to explain than outcomes on individuals,firms etc. which is often the nature of cross-sectional data.
- Another reason is that R^2 for time series regressions can be high when the dependent variable (y) is trending.
- Remember that R^2 is a measure of how large the error variance is relative to the variance of y.
- When y_t is trending, there is no problem, provided a time trend is included in the regression.
- However, when $E(y_t)$ follows a linear time trend , $\frac{SST}{n-1}$ is no longer an unbiased or consistent estimator of $Var(y_t)$.
- When y has a time trend, see pp.366-367 in your textbook for the calculation of R^2.

Seasonality

- If a time series is observed at monthly, quarterly intervals (even weekly or daily), it may exhibit **seasonality**.
- For example, weather patterns changing with seasons, holiday weeks (Christmas Effect for December) etc. may create some kind of systematic seasonal structures in some time series variables.
- The time series that do display **seasonal patterns** are often **seasonally adjusted**.
- If we study with seasonally unadjusted raw data, we can include a set of **seasonal dummy variables** to account for seasonality in the dependent variable and the independent variables, or both.

Seasonality

- When estimating a regression model using monthly data, we include 11 (not 12. Why?) seasonal dummy variables indicating whether time period t corresponds to the appropriate month.
- In this framework, January is the base month, and β_0 is the intercept for January.

$$\begin{aligned} y_t &= \beta_0 + \delta_1 feb_t + \delta_2 mar_t + ... + \delta_{11} dec_t \\ &+ \beta_1 x_{t1} + ... + \beta_k x_{tk} + u_t \end{aligned}$$

- If there is no seasonality in y_t, once the x_{tj} have been controlled for, then δ_1 through δ_{11} are all zero (the null hypothesis). This is easily tested via an F test. If the calculated F statistic is greater than the critical value, it indicates seasonality in the time series of interest.

57

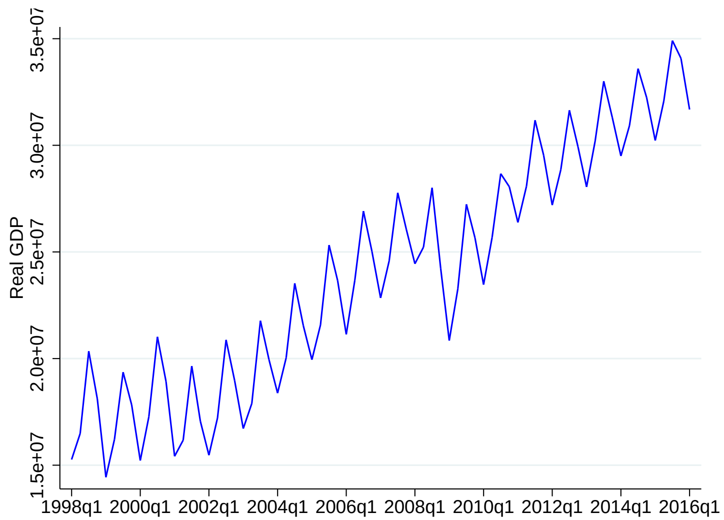
Seasonality

- If the data are quarterly, then we would include dummy variables for 3 of the four quarters, with the omitted category being the base quarter.
- Just as including a time trend in a regression has the interpretation of initially **detrending** the data, including seasonal dummies in a regression can be interpreted as **deseasonalizing** the data.

58

Seasonality Example: Turkish Quarterly Real GDP

Turkish Real GDP No Seasonal Adjustment



59

Seasonality Example: Turkish Quarterly Real GDP

Create four quarterly dummy variables; and then regress log(RGDP) on trend and three quarter dummies, quarter one will be the base group. in equation form:

$$log(\widehat{RealGDP}) = 16.48 + 0.01\ trend + 0.07\ Q_2 + 0.20\ Q_3 + 0.11\ Q_4$$

60

Seasonality Example: Turkish Quarterly Real GDP

Residuals