Introduction to Time Series

Time Series Introduction

- Time series econometrics is a **big** deal for financial econometricians and is used to:
  - Forecast
  - Model risk
  - Estimate dynamic causal effects

- These methods are heavily used by banks, stock brokers, computer scientists, engineers, etc.

- Time series questions are often predictive rather than causal

- Time series econometrics is math-intensive because it brings up some *serious* econometric issues
  - Time lags
  - Autocorrelation so $Cov(x_t, x_{t-1}) \neq 0$
  - Standard errors

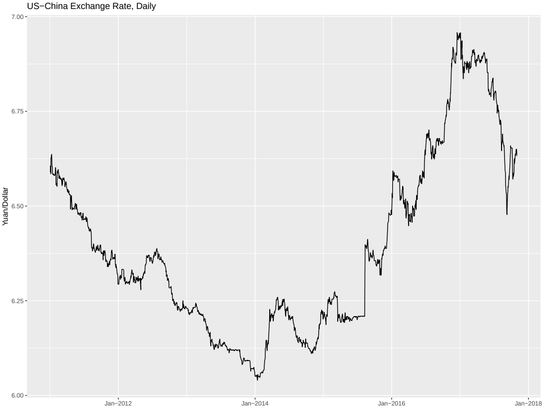
Time Series Definitions

- **Time series process** is sequence of random variables that depends on time
  - Think of a time series process as a sequence of random variables $\{X_1, X_2, ..., X_T, ...\}$ that might continue forever
  - Example: A random walk

$$X_t = X_{t-1} + e_t, \quad e_t \sim \mathcal{N}(0, 1)$$

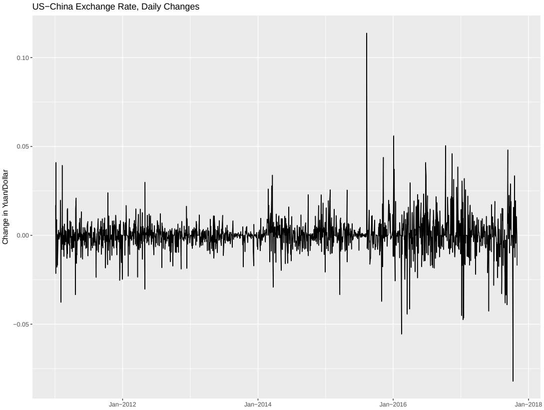
- **Time Series Data** is a realization from a particular time series process
  - Now just think of a sequence of data points, $x_1, x_2, ..., x_T$
  - Example: stock prices, unemployment, etc.
- **Time Series Econometrics**: Use our econometric tools (multiple regression, etc.) to estimate the time series process that generated time series data
  - Example: What is the statistical process underlying the stock market?

Examples of Time Series: US China Exchange Rate (levels)



- This presents a time series in **levels**

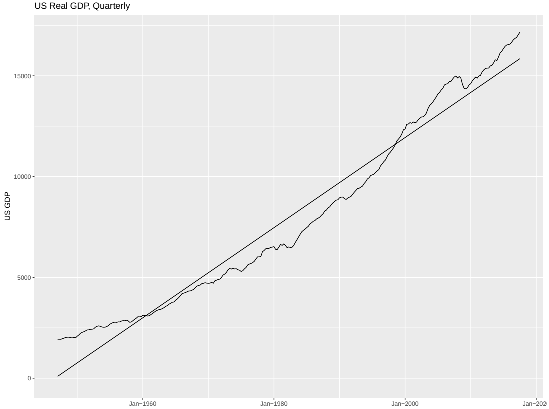
Examples of Time Series: US China Exchange Rate (changes)



- This presents the same time series in **first differences**

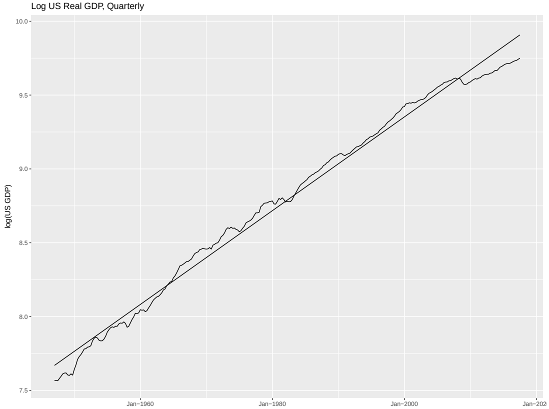
$$\Delta x_t = x_t - x_{t-1}$$

Examples of Time Series: US GDP (levels)



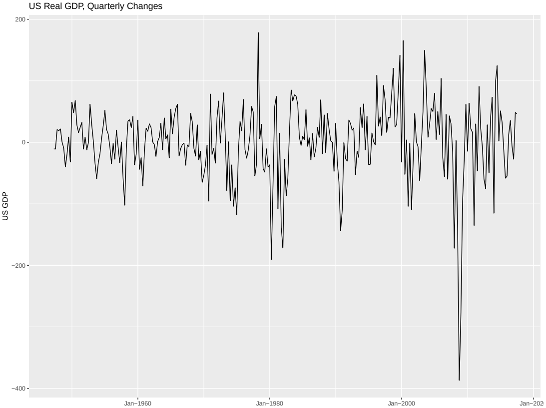
- This presents the time series of GDP in **levels**

Examples of Time Series: US GDP (logs)



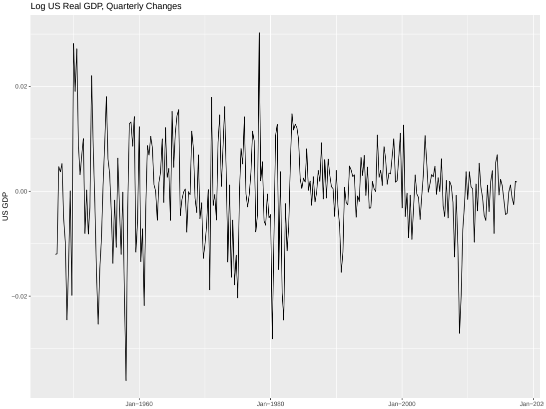
- This presents the time series of GDP in **logs**

Examples of Time Series: US GDP (changes)



- This presents the time series of GDP in **first differences**

Examples of Time Series: US GDP (log-changes)



- This presents the time series of GDP in **first differences of logs**

Time Series Transformations

- **Point of Figures:** Many ways of thinking about same time series process
  - Can use levels or logs to model evolution of x_t
  - Can use **first differences** in levels or logs to understand growth rates, $x_t - x_{t-1}$
    - When looking at growth rates, logs allow for percent changes
- Levels vs. logs vs. differences are different cuts of same process
  - Different cuts can highlight different properties – great recession looks very different in logs vs. levels
  - Which cut is right cut depends on question at hand
  - Sometimes an economic model tells us precise functional form (and sometimes not...)
  - Sometimes things only "work" in one form

Time Series Transformations

- **Point of Figures:** Many ways of thinking about same time series process
  - Can use levels or logs to model evolution of x_t
  - Can use **first differences** in levels or logs to understand growth rates, $x_t - x_{t-1}$
    - When looking at growth rates, logs allow for percent changes
- Levels versus logs versus differences are different cuts of the same process
  - Different cuts can highlight different properties – great recession looks very different in logs vs. levels
  - Which cut is right cut depends on question at hand
  - Sometimes an economic model tells us precise functional form (and sometimes not...)
  - Sometimes things will only "work" in one form

Forecasting

- As stated, time series is often more about forecasting than causal effects
- We care about very different things in forecasting than in causal effects
    - To forecast whether you will watch "Arrested Development" on Netflix no need for causality
    - Just see what other shows you watch and forecast whether you will watch "Arrested Development"
- For forecasting:
    - R^2 matters a lot!
    - We do **not** care about omitted variable bias
    - Do **not** care about causality
    - External validity is key: correlations are okay *as long as they are stable into future*

Time Series Data and Terminology

- Data set: $\{Y_1, Y_2, \ldots, Y_T\}$
  - Y_t: value of Y in period t
  - Have T observations of Y
- In this course, only consider consecutive, evenly-spaced observations
  - e.g., monthly unemployment rate from 1960-1999
- Past values of Y_t are called **lags**:
  - Y_{t-1} is the **first lag**
  - Y_{t-j} is the j^{th} lag
- Changes in Y_t are called **differences**:
  - **First difference** is $Y_t - Y_{t-1}$ and denoted by ΔY_t
- Differences often taken in logs
  - **Log first difference**: $\Delta \log Y_t = \log(Y_t) - \log(Y_{t-1})$
  - $100 \times \Delta \log(Y_t) \approx$ percent change in Y_t

Motivating Time Series: Static Models

Static Model:

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

- Simplest time series model is **static model**
  - There is a relationship between x_t and y_t over time
  - Example: the Phillips Curve (tradeoff between inflation and unemployment)
- When can we use standard OLS?
  - Strict exogeneity: $E(u_t|x_1, ..., x_T) = 0$
    - No feedback from u to future x
  - No serial correlation in errors (iid assumption)
    - $Cov(u_t, u_s) = 0$ for any $t \neq s$
- Are the OLS assumptions likely to hold?
  - Is unemployment serially uncorrelated?

Motivating Time Series 2: Finite Distributed Lag Models

Dynamic Effects:

$$y_t = \beta_0 + \delta_0 x_t + \delta_1 x_{t-1} + \delta_2 x_{t-2} + u_t$$

- **Finite distributed lag** model allows for lagged values of x to affect y
  - Example: change in technology might affect GDP slowly because of length of roll-out
- When can we use standard OLS procedures?
  - Strict exogeneity: $E(u_t | x_1, ..., x_T) = 0$
  - No serial correlation in errors (iid assumption)
    - So we must believe that x affects y slowly but *nothing else*
- Clearly we need to do better than try to shove this into standard OLS
  - We begin with modeling only *one* process

An Example Time Series

The Unemployment Rate

Quarter	Unemployment Rate	First Lag	First Difference
2016 Q3	4.9	-	-
2016 Q4	4.7	4.9	-.2
2017 Q1	4.7	4.7	0
2017 Q2	4.4	4.7	-.3
2017 Q3	4.3	4.4	.1

- First lag is number from the row above
- First difference is the difference between unemployment rate and its lag
- Obviously cannot have first lag or difference for $t = 1$

Basic Time Series Properties: Autocovariance

- Data is **serially correlated** if the value of X_t today depends on a past value of X_t
- The **first autocovariance** of X_t is defined as:

$$Cov(X_t, X_{t-1})$$

  - This can be extended to j^{th} autocovariance: $Cov(X_t, X_{t-j})$
- Estimate j^{th} autocovariance (γ_j) with:

$$\hat{\gamma}_j = \frac{1}{T} \sum_{t=j+1}^{T} (X_t - \bar{X}_{j+1,T})(X_{t-j} - \bar{X}_{1,T-j})$$

- where $\bar{X}_{j+1,T} = \frac{1}{T-j} \sum_{t=j+1}^{T} X_t$ and $\bar{X}_{1,T-j} = \frac{1}{T-j} \sum_{t=1}^{T-j} X_t$
  - Begin at $j+1$ because first j lags have no partner (i.e., there is no X_{-j})
  - Divide by T even though there are only $T-j$ terms (weird time series thing; only matters in finite samples)

Basic Time Series Properties: Autocorrelation

- Autocorrelation just normalizes autocovariance so that it lies between -1 and 1
  - Just as correlation does that for covariance
- The **first autocorrelation** of X_t is defined as:

$$\frac{Cov(X_t, X_{t-1})}{\sqrt{Var(X_t)Var(X_{t-1})}}$$

  - This can be extended to j^{th} autocorrelation:

$$\frac{Cov(X_t, X_{t-j})}{\sqrt{Var(X_t)Var(X_{t-j})}}$$

Levels and Differences Can Have Different Autocovariances

- A **random walk** is a process of the form:

$$X_t = X_{t-1} + U_t$$

  where U_t is iid over time and independent of X_{t-1}

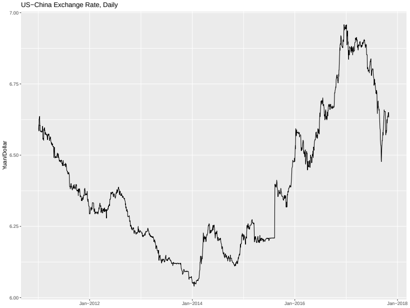
- Clearly X_t and X_{t-1} are autocorrelated:

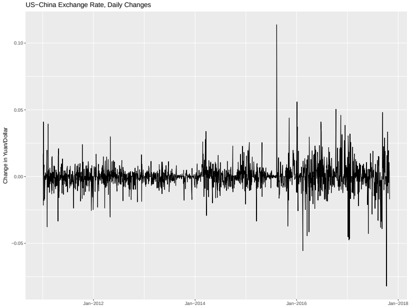
$$Cov(X_t, X_{t-1}) = Cov(X_{t-1} + U_t, X_{t-1}) = Var(X_{t-1})$$

- But what about their first-difference?
  - Under random walk, first-differences are NOT autocorrelated!

$$Cov(\Delta X_t, \Delta X_{t-1}) = Cov(X_t - X_{t-1}, X_{t-1} - X_{t-2})$$
$$= Cov(X_{t-1} + U_t - X_{t-1}, X_{t-2} + U_{t-1} - X_{t-2}) = Cov(U_t, U_{t-1}) = 0$$

The Exchange Rate Example Again



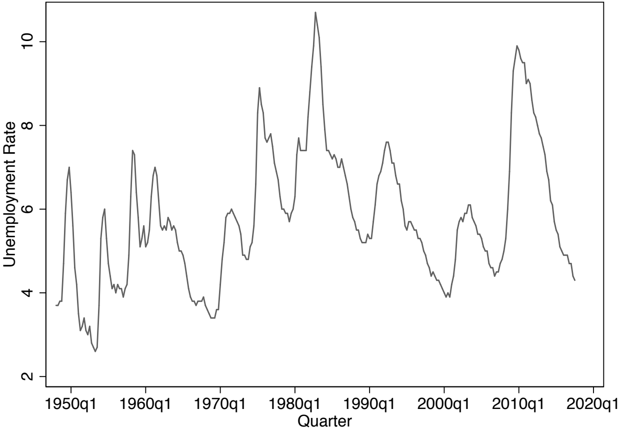


- Levels look autocorrelated
- Changes do not look autocorrelated
- *Later:* How do we test for autocorrelation?

Stationarity

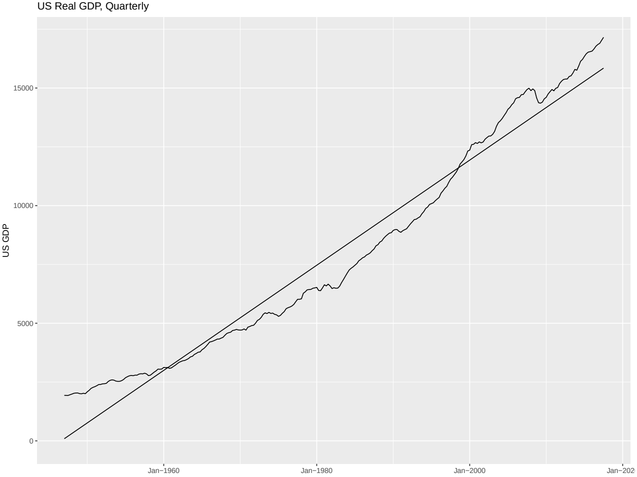
- A time series process is **stationary** if the joint distribution of $\{x_t, x_{t+1}, ..., x_{t+k}\}$ for any k does *not* depend on t
  - In words: the process has the same time series properties whether we start measuring at $t = 1$ or $t = 2$
- A time series is **stationary** if:
  - Constant mean: $E(X_t) = \mu_x$
  - Constant variance: $Var(X_t) = \sigma_x^2$
  - Autocovariances only depend on j: $Cov(X_t, X_{t-j}) = \gamma_j$

Example Stationary Process



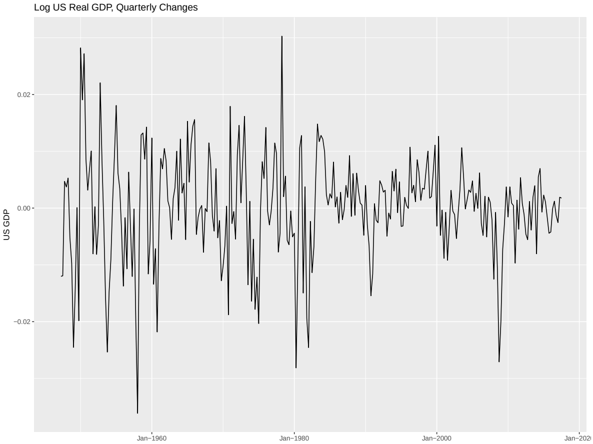
- Clearly autocorrelated and complicated process
- However— no slice of time looks like it comes from a different process

Is GDP Stationary?



- No! Mean is clearly not constant

What About Changes in GDP?



- Hard to tell
- It looks like the the variance is larger early on

Weak Dependence

- In analysis focused on causal analysis, main concern is first OLS assumption ($E(u|X) = 0$) is violated
- In time series, we do not worry about this!
- Instead worry about iid assumption
  - If Y_t is correlated with its own history then clearly it cannot be drawn independently over time!
- **Weak dependence** is the idea that x_t and x_{t+k} behave *as if* they are independent as $k \to \infty$
  - Formalizing this is far beyond the scope of this class
  - Simpler definition for this class:

$$\lim_{k \to \infty} Cov(x_t, x_{t+k}) = 0$$

  - Weak dependence replaces iid assumption

Moving Average Processes

Model:

$$x_t = e_t + \alpha_1 e_{t-1}$$

- x_t is a weighted average of two other random variables, e that occurred at different times
- **Assumption:** e_t is iid with mean 0 and variance σ_e^2
- This is called a **Moving Average Process of Order 1** or MA(1)
- We can extend this process to an MA(q) by adding more terms:

$$x_t = e_t + \alpha_1 e_{t-1} + \alpha_2 e_{t-2} + ... + \alpha_q e_{t-q}$$

- Let's derive properties of MA(1)

Properties of MA(1)

- **Question:** Find the mean and variance of the following MA(1) model where e_t is iid with zero mean and variance of σ_e^2

$$x_t = e_t + \alpha_1 e_{t-1}$$

- **Question:** Is this MA(1) model autocorrelated? If so, what is the autocovariance?
- **Question** Is MA(1) weakly dependent?

Workspace

Workspace

Summarizing MA(1)

- MA(1) is stationary:
  - The mean, variance and autocovariances all exist and do not depend on t
    - e.g., no t subscript in any of our answers
- MA(1) is weakly dependent
  - Since $Cov(X_t, X_{t-k}) = 0$ for $k > 1$
- So how do we estimate $MA(1)$?
  - Here things get tricky
  - We do not observe e so we cannot run a regression!
  - If we assume that $e \sim \mathcal{N}(0, \sigma_e^2)$ we can use method called **maximum likelihood** and if not we can use **method of moments**
- Now we move to more intuitive, but more complicated model called AR(p)

Autoregression

- For now, assume stationarity
- **Autoregression** is a forecasting model which uses past values of Y (that is, $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}$) to forecast Y_t
  - In autoregression we regress Y_t against its own lagged values
- Number of lags used as regressors is called the *order of the autoregression*
  - First order autoregression: regress Y_t on Y_{t-1}
  - p^{th} order autoregression: regress Y_t on $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}$
- Notation: $AR(p)$ is p^{th} order autoregressive model

First Order Autoregressive (AR(1)) Model

AR(1) Model:

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

- y_t is direct function of its past self and a noise term
  - We assume that u_t is mean 0 and iid over time (hence independent of y_{t-1})
- **Key Assumption:** AR(1) process is **stable** if $|\rho| < 1$
  - Necessary condition for weak dependence and stationarity
- This can be extended to autoregressive process of order p, $AR(p)$:

$$y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + ... + \rho_p y_{t-p} + u_t$$

- What are properties of AR(1) process?

Properties of AR(1)

- **Question:** Find the mean and variance of the following AR(1) model where u_t is iid with zero mean and variance of σ_u^2. Be clear where you impose stationarity.

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

- **Question:** Is this AR(1) model autocorrelated? If so, what is the autocovariance?
- **Question** Is AR(1) weakly dependent?

Workspace

Workspace

Summarizing AR(1)

- AR(1) is only stationary if $|\rho| < 1$
  - Things go haywire if this isn't true
- AR(1) is more complicated than the MA(1):
  - Autocovariances non-zero forever, although die out eventually if $|\rho| < 1$
  - AR(1) thus weakly dependent if $|\rho| < 1$
- Why use AR(1)?
  - Most commonly modeled time series process
  - Often very realistic process
    - Processes tend to persist over long stretches, unlike MA(1) where processes are not correlated for one period then suddenly uncorrelated
    - Many economic models predict AR(1) processes and random walks
  - Unlike MA(1), can be estimated with OLS!

Estimating AR(1) Model

- The AR(1) model is estimated with OLS
- Assumptions *for consistency* (no need for unbiasedness):
  1. Contemporaneous Exogeneity: $E(u_t|y_t) = 0$
     - Same as the original OLS assumption
     - We CANNOT have strict exogeneity (why?)
  2. Stationarity and Weak Dependence
     - This is a substantially weaker assumption than iid
  3. No perfect multicollinearity
- Additional assumptions *for inference*
  1. Homoscedasticity: $Var(u|Y) = \sigma_Y^2$
  2. No serial correlation in u: $E(u_s u_t|y_t, y_s) = 0$
- With these additional assumptions, for large T, all usual testing procedures will be valid

Example AR(1): GDP Growth Rates

How good of a predictor is lagged GDP growth?

$$\Delta \log GDP_t = \beta_0 + \beta_1 \Delta \log GDP_{t-1} + U_t$$

Results:

$$\widehat{\Delta \log GDP_t} = \underset{(.001)}{.005} + \underset{(.056)}{.372} \times \Delta \log GDP_{t-1}$$

- This is *not* a causal relationship
- What "causes" GDP growth (positive or negative) to be correlated over time?
  - Roll out of new inventions
  - Upward trend in education
  - Bank failures might take time to cascade through a network
- Putting lagged GDP on the RHS captures all of these relationships