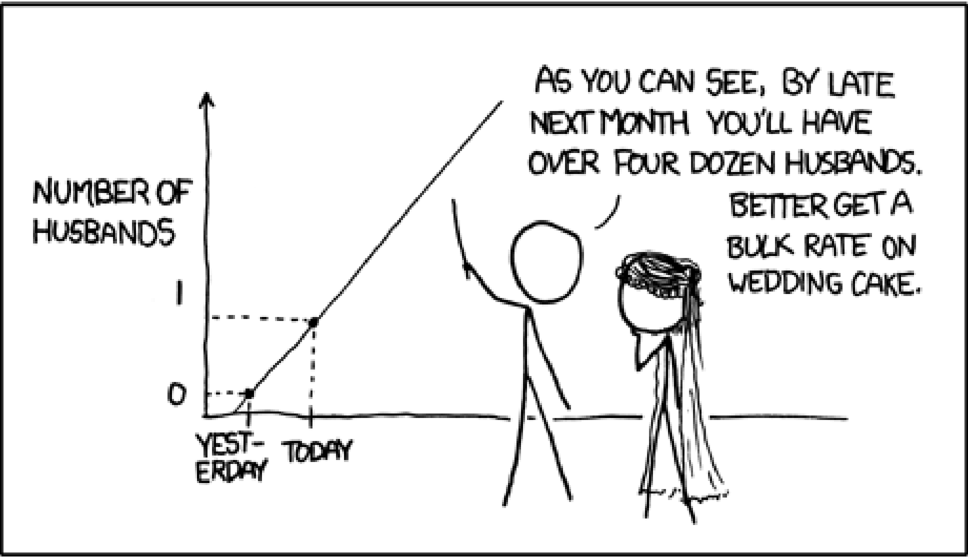
MY HOBBY: EXTRAPOLATING



Recitation 10: Time series econometrics

14.32 Fall 2023

Ian Sapollnik

Dec 8, 2023

What is time series data and why do we care?

Time series data is collected on a **single** entity over many periods of time. We need to exercise extra caution because in general such data will **not** be i.i.d.

As a result, many of the tools and frameworks we have developed so far might fail since they require i.i.d data.

We say that our data is observed for T periods, and an observation at a specific unit in time is Y_t

Some definitions

- Y_{t-j}: the jth lag of Y_t
- Stationarity: $\{Y_t : t = \dots, -1, 0, 1, \dots\}$ is stationary if $\{Y_t, \dots, Y_{t+k}\}$ has the same distribution for all k.
- Autocovariance $\gamma_k = \text{Cov}\,(Y_t, Y_{t+k})$ depends only on k if sequence is stationary.

Autoregressive (AR) model

An autoregressive $AR\,(r)$ model takes the form

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \cdots + \beta_r Y_{t-r} + \varepsilon_t$$

This is not a *causal* model, we use such a model for forecasting or prediction. β_k describes how good Y_{t-k} is at predicting Y_t.
Question: do we are about

- Endogeneity?

Autoregressive (AR) model

An autoregressive $AR(r)$ model takes the form

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \cdots + \beta_r Y_{t-r} + \varepsilon_t$$

This is not a *causal* model, we use such a model for forecasting or prediction. β_k describes how good Y_{t-k} is at predicting Y_t.
Question: do we are about

- Endogeneity? No!
- Omitted variable bias?

Autoregressive (AR) model

An autoregressive $AR\,(r)$ model takes the form

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \cdots + \beta_r Y_{t-r} + \varepsilon_t$$

This is not a *causal* model, we use such a model for forecasting or prediction. β_k describes how good Y_{t-k} is at predicting Y_t.
Question: do we are about

- Endogeneity? No!
- Omitted variable bias? No!
- Standard errors/significance?

Autoregressive (AR) model

An autoregressive $AR\,(r)$ model takes the form

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \cdots + \beta_r Y_{t-r} + \varepsilon_t$$

This is not a *causal* model, we use such a model for forecasting or prediction. β_k describes how good Y_{t-k} is at predicting Y_t.
Question: do we are about

- Endogeneity? No!
- Omitted variable bias? No!
- Standard errors/significance? Kind of?
- R^2?

Autoregressive (AR) model

An autoregressive $AR(r)$ model takes the form

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \cdots + \beta_r Y_{t-r} + \varepsilon_t$$

This is not a *causal* model, we use such a model for forecasting or prediction. β_k describes how good Y_{t-k} is at predicting Y_t.
Question: do we are about

- Endogeneity? No!
- Omitted variable bias? No!
- Standard errors/significance? Kind of?
- R^2? Yes!

Lag selection in an $AR(r)$ model

We don't want to pick a very large lag r for our model because:

- we have finite data
- very distant lags are unlikely to be predictive

We can pick r through *information criterion* such as the Bayesian (BIC).

That is, pick the r that minimizes

$$BIC(r) = \ln\left(\frac{SSR(r)}{T}\right) + (r+1)\frac{\ln(T)}{T}$$

Autoregressive distributed lag model

We can also add controls to our model in an autoregressive distributed lag model (ADL) of the form

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \cdots + \beta_r Y_{t-r} + \delta_0 X_t + \delta_1 X_{t-1} + \cdots + \delta_p X_{t-p} + \varepsilon_t$$

We continue to not have a causal interpretation, focusing only on prediction and forecasting.

We can check if X has predictive power for Y by testing $H_0 : \delta_1 = \cdots = \delta_p = 0$ with an F-statistic.

We can again use BIC to pick lag length, and p and r are allowed to differ.

Adding causality back into the picture

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \cdots + \beta_r Y_{t-r} + \delta_0 X_t + \delta_1 X_{t-1} + \cdots + \delta_p X_{t-p} + \varepsilon_t$$

To think about causation again, we need to make some assumptions

- Exogeneity: $\mathbb{E}\left[\varepsilon_t \mid X_t, X_{t-1}, \ldots\right] = 0$
- Stationarity: (X_t, Y_t) is stationary
- Independence: (X_t, Y_t) and (X_{t-k}, Y_{t-k}) become more uncorrelated as k grows
- No outliers or multicolinearity, as before

Interpretting dynamic causal effects

$$Y_t = \beta_0 + \delta_0 X_t + \delta_1 X_{t-1} + \cdots + \delta_p X_{t-p} + \varepsilon_t$$

In this model, we can interpret

- δ_k as the contemporaneous effect of X on Y
- δ_k as a k-period dynamic multiplier, holding X_{-k} constant
- $\delta_0 + \delta_1 + \cdots + \delta_k$ is a cumulative dynamic multiplier

Standard errors in time series data

Often, ε is autocorrelated, since the omitted factors included in ε can themselves be serially correlated. This autocorrelation:

- does not affect the consistency of OLS
- does not introduce bias
- **does** affect how we should calculate standard errors

We need to rely on heteroskedasticity and auto-correlation (HAC) robust standard errors. The Newey-West standard error is the most common of these.

In general, cannot allow auto-correlation to exist for an arbitrarily large amount of time, as CLT will not hold.

Please fill out your subject evaluations!

`https://eduapps.mit.edu/`
`subjeval/studenthome.htm`