Financial Econometrics Workshop 2

Binzhi Chen

August 23, 2020

1 Introduction

In this workshop we will introduce basic time series model in STATA. Date transformation and model selection approach will also be introduced in this workshop.

2 Input Data and Date Transformation

The first step in STATA is input data, make sure your .dta file is in the right working directory. In this workshop we will use the dataset: monthly JP Morgan/Goldman Sachs stock price data from 2000/01/01 to 2014/12/01.

Menu:
File → Change Working Directory → STATA Workshop

Command:

```
use gs.dta,clear
generate date_1 = date(date,"YMD") //transform date as year-month-day
generate datem = mofd(date_1) //the monthly date (months since 1960m1) containing date_1
format datem %tm // %tm encoded monthly date (months since 1960m1)
tsset datem
```

Note:

`tsset` declares data to be time-series data.

3 Graphic Analysis and Data Transformation

In this section, you will learn how to combine two dependent variables in the same graph

Command:

```
graph twoway line gs_adjclose datem, yaxis(1) sort || line
    jpm_adjclose datem, yaxis(2) sort
graph twoway connected gs_adjclose datem,yaxis(1)|| connected
    jpm_adjclose datem, yaxis(2) sort msymbol(Th)
dfuller gs_adjclose // DF test for original time series
dfuller jpm_adjclose // DF test for original time series

g log_gs_adjclose = log(gs_adjclose)

g dlog_gs_adjclose = d.log_gs_adjclose

dfuller dlog_gs_adjclose // DF test for dlog time series

g log_jpm_adjclose = log(jpm_adjclose)

g dlog_jpm_adjclose = d.log_jpm_adjclose

dfuller dlog_jpm_adjclose // DF test for dlog time series

// We use jpm data as an example for autocorrelation function and
    partial autocorrelation function
corrgram dlog_jpm_adjclose,lags(10)

ac dlog_jpm_adjclose,lags(10) //autocorrelation function
pac dlog_jpm_adjclose,lags(10) //partial autocorrelation function
```

Note:

We will introduce DF test in next workshop.

Note:

By log-differencing, we can transform non-stationary time series data into stationary time series data.

4 ARIMA Model Selection

As a benchmark model, $ARIMA(p, d, q)$ can be divided by three parts:
$AR(p)$, Integrated(d) and $MA(q)$.

Command:

```
// We use gs data as an example for ARIMA
arima dlog_gs_adjclose,arima(1,0,0)
estat ic //AIC BIC
est store m1 //store regression result in m1

quietly arima dlog_gs_adjclose,arima(2,0,0)
est store m2

quietly arima dlog_gs_adjclose,arima(0,0,1)
est store m3

quietly arima dlog_gs_adjclose,arima(0,0,2)
est store m4

quietly arima dlog_gs_adjclose,arima(1,0,1)
est store m5

quietly arima dlog_gs_adjclose,arima(1,0,2)
est store m6

quietly arima dlog_gs_adjclose,arima(2,0,1)
est store m7

quietly arima dlog_gs_adjclose,arima(2,0,2)
est store m8

est table m1 m2 m3 m4 m5 m6 m7 m8,stats(aic bic)
```

Note:

For demonstration, the lag length in ARIMA <= 2.

Note:

`quietlyly` command does not show the regression result.

Note:

We can also use `for` loop to select model automatically.