Problem Set 11

Due Wednesday April 9, 4pm

Data Exercises

(1) Use the file "realgdpgrowth.xlsx". It includes the following quarterly variables

- t3month = rate on 3-month T-bill (TB3MS)
- t1year = rate on 1-year T-bill (GS1)
- t5year = rate on 5-year T-bill (GS5)
- t10year = rate on 10-year T-bill (GS10)
- AAA = rate on AAA corporate bonds (AAA)
- BAA = rate on BAA corporate bonds (BAA)
- gdp = real GDP (percent change from preceding period)

Note: The variables t1year, t5year and t10year are missing until 1953q2.

(a) Create the transformed variables

(i) spread1 = t1year-t3month

(ii) spread2 = t10year-t3month

(iii) corporate = BAA-AAA

(iv) dt3 = t3month-L.t3month

(v) dt12 = t1year-L.t1year

Describe in words the variables you created.

(b) Test the following hypotheses. For each, use three lags of all variables, and restrict the sample to 1954q2-2024q3 so all have the same number of observations.

(i) *dt3* does not Granger-cause *gdp*

(ii) *dt12* does not Granger-cause *gdp*

(iii) *spread1* does not Granger-cause *gdp*

(iv) *spread2* does not Granger-cause *gdp*

(v) *corporate* does not Granger-cause *gdp*

Interpret your findings.

(c) Of these five, which would you select to forecast GDP? Explain your reasoning.

(d) Use your selected model to make point and interval forecast for GDP for 2024Q4, 2025Q1-Q3.

(2) Check whether returns on the highest rated (AAA) and lower rated (BAA) bonds have unit roots. Use monthly data from FRED (labels: AAA and BAA). Then check whether the high-yield (junk) bond spread has a unit root. Interpret your findings.

Theoretical Questions

(3) Consider AR(3)

$$y_t = \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \varepsilon_t$$

and let $\Delta y_t := y_t - y_{t-1}$.

(a) Rewrite the above AR(3) as

$$\Delta y_t = \mu + (\beta_1 + \beta_2 + \beta_3 - 1) y_{t-1} + \gamma_1 \Delta y_{t-1} + \gamma_2 \Delta y_{t-2} + \varepsilon_t \tag{1}$$

and express μ, γ_1, γ_2 in terms of $\alpha, \beta_1, \beta_2, \beta_3$.

(b) Suppose that y_t has a unit root. What does it imply about the coefficients in the regression (1)?

(4) Let $\varepsilon_t \sim$ i.i.d.$(0, \sigma^2)$, $u_t \sim$ i.i.d.$(0, \sigma_u^2)$, $v_t \sim$ i.i.d.$(0, \sigma_v^2)$. Assume that $\varepsilon_t, u_s, v_\tau$ are independent from each other for any t, s, τ. Let

$$x_t = \sum_{i=1}^{t} \varepsilon_i + u_t, \quad y_t = 0.4 \sum_{i=1}^{t} \varepsilon_i + v_t.$$

(a) Calculate the mean and variance of x_t and y_t. Is x_t stationary? Is y_t stationary?

(b) Is $\Delta x_t := x_t - x_{t-1}$ stationary? Is $\Delta y_t := y_t - y_{t-1}$ stationary? (Calculate their mean, variance and autocovariances.)

(c) Guess the value of θ such that $y_t - \theta x_t$ is stationary. Are x_t and y_t cointegrated?

(d) Calculate the mean, variance and autocovariances of $y_t - \theta x_t$ for the value of θ from part (c). Are your results consistent with stationarity?

(e) (For this part you will need to use Matlab or any other software) Simulate x_t and y_t for $t = 1, \ldots, 500$ from $\varepsilon_t, u_t, v_t \sim$ i.i.d. $\mathcal{N}(0, 1)$. Regress y_t on a constant and x_t. Discuss the estimated coefficients and compare with spurious regression.

4. (a). ① $E[X_t] = E\left[\sum_{i=1}^{t} \varepsilon_i + u_t\right] = \sum_{i=1}^{t} E[\varepsilon_i] + E[u_t] = 0 + 0 = 0$,

$Var(X_t) = Var\left(\sum_{i=1}^{t} \varepsilon_i + u_t\right) = Var\left(\sum_{i=1}^{t} \varepsilon_i\right) + Var(u_t) + 0$ since ε_i and u_i all follows their i.i.d, so $\sum \varepsilon_i$ and u_i are uncorrelated, And since ε_i are i.i.d $\forall i$,

$Var(X_t) = Var\left(\sum_{i=1}^{t} \varepsilon_i\right) + Var(u_t) = \sum_{i=1}^{t} Var(\varepsilon_i) + Var(u_t) = t \cdot \sigma^2 + \sigma_u^2$

$\forall h \neq 0$. WLOG, consider $h > 0$.

$Cov(X_t, X_{t+h}) = Cov\left(\sum_{i=1}^{t} \varepsilon_i + u_t, \sum_{i=1}^{t+h} \varepsilon_i + u_{t+h}\right) = Cov\left(\sum^{t} \varepsilon_i, \sum^{t+h} \varepsilon_i\right) + Cov(u_t, u_{t+h})$
$+ Cov\left(\sum^{t} \varepsilon_i, u_{t+h}\right) + Cov\left(u_t, \sum^{t+h} \varepsilon_i\right) = Cov\left(\sum^{t} \varepsilon_i, \sum^{t+h} \varepsilon_i\right) + 0 = Cov\left(\sum^{t} \varepsilon_i, \sum^{t} \varepsilon_i\right) +$
$Cov\left(\sum^{t} \varepsilon_i, \sum_{i=t+1}^{t+h} \varepsilon_i\right) = \sum^{t} Var(\varepsilon_i) + 0 = t\sigma^2$

The $Var(X_t)$ and $\rho(t,h)$ depends on t, So X_t is not stationary.

---

② $E[Y_t] = E\left[0.4\sum_{i=1}^{t} \varepsilon_i + v_i\right] = 0.4\sum_{i=1}^{t} E[\varepsilon_i] + E[v_i] = 0 + 0 = 0$,

$Var(Y_t) = Var\left(0.4\sum_{i=1}^{t} \varepsilon_i + v_i\right) = 0.16 \cdot Var\left(\sum_{i=1}^{t} \varepsilon_i\right) + Var(v_i) + 0$ since ε_i and v_i all follows their i.i.d, so $\sum \varepsilon_i$ and v_i are uncorrelated, And since ε_i are i.i.d $\forall i$,

$Var(Y_t) = 0.16\, Var\left(\sum_{i=1}^{t} \varepsilon_i\right) + Var(v_i) = 0.16 \cdot \sum_{i=1}^{t} Var(\varepsilon_i) + Var(v_i) = 0.16\, t \cdot \sigma^2 + \sigma_v^2$

The $Var(Y_t)$ depends on t, So it's enough to conclude that Y_t is not stationary.

(b). ① $\Delta x_t = x_t - x_{t-1} = \sum^{t} \varepsilon_i + u_t - \sum^{t-1} \varepsilon_i + u_{t-1} = \varepsilon_t + u_t - u_{t-1}$.

mean: $E[\Delta x_t] = E[\varepsilon_t + u_t - u_{t-1}] = E[\varepsilon_t] + E[u_t] - E[u_{t-1}] = 0$.

Var: $Var(\Delta x_t) = Var(\varepsilon_t + u_t - u_{t-1}) = Var(\varepsilon_t) + Var(u_t) + Var(u_{t-1})$ since they're i.i.d

$= \sigma^2 + \sigma_u^2 + \sigma_u^2 = \sigma^2 + 2\sigma_u^2$

AC: $\gamma(t,h), = Cov(\Delta x_t, \Delta x_{t+h}) = Cov(\varepsilon_t + u_t - u_{t-1}, \varepsilon_{t+h} + u_{t+h} - u_{t+h-1})$

$= Cov(\varepsilon_t, \varepsilon_{t+h}) + 0 + 0 + 0 + Cov(u_t, u_{t+h}) + Cov(u_t, -u_{t+h-1}) + 0 + Cov(-u_{t-1}, u_{t+h}) + Cov(-u_{t-1}, -u_{t+h-1})$

$$= \begin{cases} \sigma^2 + \sigma_u^2 + \sigma_u^2 & \text{if } h = 0, \\ -\sigma_u^2 & \text{if } h = +1 \\ 0 & \text{if } h \geq 2 \end{cases}$$

The $\gamma(t,h) = \gamma(h)$, doesn't depend on t so Δx_t is Covariance stationary, we know it's also mean stationary so Δx_t is stationary.

---

② $\Delta y_t = y_t - y_{t-1} = 0.4\sum^{t} \varepsilon_i + v_t - 0.4\sum^{t-1} \varepsilon_i + v_{t-1} = 0.4\varepsilon_t + v_t - v_{t-1}$

mean: $E[\Delta y_t] = E[0.4\varepsilon_t + v_t - v_{t-1}] = 0$.

Var: $Var(\Delta y_t) = Var(0.4\varepsilon_t + v_t - v_{t-1}) = 0.16\,Var(\varepsilon_t) + Var(v_t) + Var(v_{t-1})$ since they're i.i.d

$= 0.16\sigma^2 + 2\sigma_v^2$

AC: Similarly, we have $\gamma(t,h) = cov(y_t, y_{t+h})$

$$= \begin{cases} 0.16\sigma^2 + 2\sigma_v^2 & \text{if } h = 0, \\ -\sigma_v^2 & \text{if } h = +1 \\ 0 & \text{if } h \geq 2 \end{cases}$$

$\gamma(t,h) = \gamma(h)$, doesn't depend on t so it's cov stationary.

cov stationary + mean stationary $\Rightarrow$ Δy_t is stationary.

(c). & (d). $Y_t - \theta X_t$, guess $\theta = 0.4$.

Then, $Y_t - 0.4X_t = 0.4\sum^{t} \varepsilon_i + V_t - \left(0.4\sum^{t}\varepsilon_i + 0.4u_t\right) = V_t - 0.4u_t \equiv A$

I would think $Y_t - 0.4X_t$ is cointegrated since it's obvious that it has constant mean, var, Autocov.

Check

mean: $E[A] = E[V_t] + E[-0.4u_t] = 0$. mean stat.

Var: $Var(A) = Var(V_t - 0.4u_t) = Var(V_t) + Var(-0.4u_t) + 0 = \sigma_v^2 - 0.16\sigma_u^2$ Var stat

AC: $\gamma(t,h) = Cov(A_t, A_{t+h}) = Cov(V_t - 0.4u_t, V_{t+h} - 0.4u_{t+h}) = \begin{cases} \sigma_v^2 - 0.16\sigma_u^2 & h=0, \\ 0 & h \geq 1 \end{cases}$

$\gamma(t,h) = \gamma(h)$. $Y_t - \theta X_t$ is stationary

Since Y_t, X_t are non-stationary (they may have unit root),

but $Y_t - \theta X_t$ is stationary for some $\theta = 0.4$, we call X, Y. cointegrated.

(e). check another file.

(3) Consider AR(3)

$$y_t = \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \varepsilon_t$$

and let $\Delta y_t := y_t - y_{t-1}$.

(a) Rewrite the above AR(3) as

(1) $$\Delta y_t = \mu + (\beta_1 + \beta_2 + \beta_3 - 1) y_{t-1} + \gamma_1 \underbrace{\Delta y_{t-1}}_{=(y_{t-1}-y_{t-2})} + \gamma_2 \underbrace{\Delta y_{t-2}}_{=(y_{t-2}-y_{t-3})} + \varepsilon_t$$

and express $\boxed{\mu, \gamma_1, \gamma_2}$ in terms of $\alpha, \beta_1, \beta_2, \beta_3$.

(b) Suppose that y_t has a unit root. What does it imply about the coefficients in the regression (1)?

3.(a). $\Delta y_t \equiv y_t - y_{t-1}$. Following AR(3), with intercept, we have

$y_t = \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \varepsilon_t$

$y_{t-1} = \alpha + \beta_1 y_{t-2} + \beta_2 y_{t-3} + \beta_3 y_{t-4} + \varepsilon_{t-1}$

$\Delta y_t = y_t - y_{t-1} = (\alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \varepsilon_t) - (\alpha + \beta_1 y_{t-2} + \beta_2 y_{t-3} + \beta_3 y_{t-4} + \varepsilon_{t-1})$

OR I should consider

$\Delta y_t = y_t - y_{t-1} = (\alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \varepsilon_t) - y_{t-1}$

$= \alpha + (\beta_1+\beta_2+\beta_3-1)\cdot y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \varepsilon_t - (\beta_2+\beta_3) y_{t-1}$

$= \alpha + (\beta_1+\beta_2+\beta_3-1)\cdot y_{t-1} + \underline{\beta_2 y_{t-2}} + \underline{\beta_3 y_{t-3}} - \beta_2 y_{t-1} - \beta_3 y_{t-1} + \varepsilon_t$

$= \alpha + (\beta_1+\beta_2+\beta_3-1)\cdot y_{t-1} + (-\beta_2 y_{t-1} - \beta_3 y_{t-1}) + (\underline{\beta_2 y_{t-2}} + \underline{\beta_3 y_{t-2}}) + (\underline{-\beta_3 y_{t-2}} + \underline{\beta_3 y_{t-3}}) + \varepsilon_t$

$= \alpha + (\beta_1+\beta_2+\beta_3-1)\cdot y_{t-1} + -(\beta_2+\beta_3)\cdot(y_{t-1} - y_{t-2}) - \beta_3 (y_{t-2} - y_{t-3}) + \varepsilon_t$

$= \mu + (\beta_1+\beta_2+\beta_3-1)\cdot y_{t-1} + \gamma_1 \cdot \Delta y_{t-1} + \gamma_2 \cdot \Delta y_{t-2} + \varepsilon_t.$

where $\mu = \alpha$ $\quad \gamma_1 = -(\beta_2+\beta_3)$ $\quad \gamma_2 = -\beta_3$

(b) $Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \beta_3 Y_{t-3} + \varepsilon_t$

$$(Y_t - \beta_1 Y_{t-1} - \beta_2 Y_{t-2} - \beta_3 Y_{t-3}) = \alpha + \varepsilon_t,$$

$$(1 - \beta_1 L - \beta_2 L^2 - \beta_3 L^3) \cdot Y_t = \alpha + \varepsilon_t,$$

$$a(L)\, Y_t = \alpha + \varepsilon_t$$

when Y_t has a unit root, we have that $a(1) = 0$, $\Rightarrow (1-\beta_1 - \beta_2 - \beta_3) = 0$

The second coefficient, $(1-\beta_1 - \beta_2 - \beta_3)$ will equal to 0 so the Y_t on RHS will vanish, Then we will have,

$$\Delta Y_t = \alpha + \gamma_1 \cdot \Delta Y_{t-1} + \gamma_2 \cdot \Delta Y_{t-2}, + \varepsilon_t,$$

$$= \alpha - (\beta_2 + \beta_3) \cdot (Y_{t-1} - Y_{t-2}) - \beta_3 (Y_{t-2} - Y_{t-3}) + \varepsilon_t$$

It means the effect of Y_{t-1} on ΔY_t cancel out, only some Y_{t-1} come from the effect of past difference ΔY_{t-1}. ΔY_t now only depend on intercept and past difference ΔY_{t-1}, ΔY_{t-2} but not Y_{t-1}.