

QE Problem Set 7

$$\begin{aligned}
 X_{St+} &= \beta_0 + \beta_1 X_{St+1} + \varepsilon_{St+} \\
 &= \beta_0 + \beta_1 (\beta_0 + \beta_1 X_{St+2} + \varepsilon_{St+1}) + \varepsilon_{St+} \\
 &= (\beta_0 + \beta_1 \beta_0) + \beta_1^2 X_{St+2} + \varepsilon_{St+} + \beta_1 \varepsilon_{St+1} \\
 &= \dots \\
 &= \beta_0 (1 + \beta_1 + \beta_1^2 + \dots) + \beta_1^t \\
 &\quad \beta_0 (1 + \beta_1 + \beta_1^2 + \dots + \beta_1^{t-1}) \\
 &\quad + \beta_1^t X_S \\
 &\quad + (\varepsilon_{St+} + \beta_1 \varepsilon_{St+1} + \beta_1^2 \varepsilon_{St+2} + \dots + \beta_1^{t-1} \varepsilon_{St+}) \\
 &= \beta_0 \sum_{i=0}^{t-1} \beta_1^i + \beta_1^t X_S + \sum_{i=0}^{t-1} \beta_1^i \varepsilon_{St+i}
 \end{aligned}$$

$$x_t = \beta_0 \sum_{i=0}^{t-1} \beta_i^i + \beta_t^t x_0 + \sum_{i=0}^{t-1} \beta_i^i \varepsilon_{t-i}$$

Given that β_0 and β_i are constants, $f_0 \sum_{i=0}^{t-1} \beta_i$ is independent of E_i for all $i \geq 1$.

Given that X_0 is independent of ε_i for all $i \geq 1$, since β_1 is a constant, $\beta_1^+ X_0$ is independent of ε_i for all $i \geq 1$.

Dependence of X_t on $\{\varepsilon_i\}$ is given entirely by the term $\sum_{i=0}^{t-1} \beta_i \varepsilon_{t-i}$. X_t depends directly only on $\{\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_1\}$. Since $\{\varepsilon_i\}$ is iid, X_t is independent of $\{\varepsilon_{t+1}, \varepsilon_{t+2}, \dots\}$.

$$\begin{aligned}
 & \mathbb{E}(X_t) \\
 &= \mathbb{E}(\beta_0 \sum_{i=0}^{t-1} \beta_i^i + \beta_i^t x_0 + \sum_{i=0}^{t-1} \beta_i^i \varepsilon_{t-i}) \\
 &= \beta_0(1 - \beta_1^t)(1 - \beta_1) + \beta_1^t \mathbb{E}(x_0) + \sum_{i=0}^{t-1} \beta_i^i \mathbb{E}(\varepsilon_{t-i}) \\
 &\text{by linearity of expectations, since } \beta_0 \text{ and } \beta_i \text{ are constants} \\
 &\Rightarrow \beta_0(1 - \beta_1^t)/(1 - \beta_1) + \beta_1^t \beta_0/(1 - \beta_1) \\
 &\text{given } \mathbb{E}(x_0) = \beta_0/(1 - \beta_1) \text{ and } \mathbb{E}(\varepsilon_i) = 0 \text{ for } i \geq 1 \\
 &= \beta_0/(1 - \beta_1)
 \end{aligned}$$

$$\begin{aligned}
 & \text{var}(x_t) \\
 &= \text{var}(\beta_0 \sum_{i=0}^{t-1} \beta_i^i + \beta_1 x_0 + \sum_{i=0}^{t-1} \beta_i^i \varepsilon_{t-i}) \\
 &= \beta_1^{2t} \text{var}(x_0) + \sum_{i=0}^{t-1} \beta_i^i \text{var}(\varepsilon_{t-i}) \\
 & \text{Since } \beta_0 \text{ and } \beta_1 \text{ are constants, } x_0 \text{ is independent} \\
 & \text{of } \varepsilon_i \text{ for } i \geq 1, \text{ and } \{\varepsilon_i\} \text{ is iid} \\
 &= \beta_1^{2t} \sigma_\varepsilon^2 / (1 - \beta_1^2) + \sum_{i=0}^{t-1} \beta_i^{2i} \sigma_\varepsilon^2 \\
 & \text{given } \text{var}(x_0) = \sigma_\varepsilon^2 / (1 - \beta_1^2) \text{ and } \text{var}(\varepsilon_i) = \sigma_\varepsilon^2 \text{ for} \\
 & \text{all } i \geq 1 \\
 &= \beta_1^{2t} \sigma_\varepsilon^2 (1 - \beta_1^2) + (1 - \beta_1^{2t}) \sigma_\varepsilon^2 / (1 - \beta_1^2) \\
 &= \sigma_\varepsilon^2 / (1 - \beta_1^2)
 \end{aligned}$$

$$\begin{aligned}
 & \text{cov}(X_S, X_{S+i}) \\
 &= \text{cov}(\beta_0 \sum_{i=0}^{t-1} \beta_i + \beta_0^T X_S + \sum_{i=0}^{t-1} \beta_i^T \varepsilon_{S-i}, \\
 & \quad \beta_0 \sum_{i=0}^{t-1} \beta_i + \beta_0^T X_S + \sum_{i=0}^{t-1} \beta_i^T \varepsilon_{S+i}) \\
 &= \text{cov}(\beta_0^T X_S + \sum_{i=0}^{t-1} \beta_i^T \varepsilon_{S-i}, \beta_0^T X_S + \sum_{i=0}^{t-1} \beta_i^T \varepsilon_{S+i}) \\
 &= \beta_0^T \text{var}(X_S) \beta_0 + \text{cov}(\varepsilon_S + \beta_0^T \varepsilon_{S+i}, \varepsilon_S + \beta_0^T \varepsilon_{S+i}) \\
 &+ \text{cov}(\varepsilon_S + \beta_0^T \varepsilon_{S+i}, \beta_0^T \varepsilon_{S+i} + \dots + \beta_1^T \varepsilon_{S+i}) \\
 \end{aligned}$$

Since X_S is independent of ε_i for $i \geq S+1$

since β_0 and β_i are constants

$$\begin{aligned}
 & \text{cov}(X_S + \varepsilon_S + \beta_0^T \varepsilon_{S+i}, \beta_0^T X_S + \sum_{i=0}^{t-1} \beta_i^T \varepsilon_{S+i}) \\
 &= \text{cov}(X_S + \varepsilon_S, \beta_0^T X_S + \varepsilon_{S+i} + \dots + \beta_1^T \varepsilon_{S+i})
 \end{aligned}$$

$$\begin{aligned}
 &= \text{cov}(X_S + \epsilon_S + \beta_1^{-1} \epsilon_{S+1}, \beta_1^T X_S + \epsilon_{S+1} + \dots + \beta_1^{T+1} \epsilon_{S+1}) \\
 &= \beta_1^T \text{var}(X_S) + \beta_1^T \text{cov}(X_S, \epsilon_S) + \beta_1^T \text{var}(\epsilon_{S+1}) \\
 &\quad \text{by bilinearity of covariance, since } X_S \text{ is independent of } \epsilon_i \text{ for } i \geq S+1 \text{ and } \{\epsilon_i\}_{i \geq 1} \text{ is iid.} \\
 &= \beta_1^T \sigma_\epsilon^2 / (1 - \beta_1^2) \cancel{+ \beta_1^T \text{var}(\epsilon_S)} \cancel{+ 2\beta_1^T \sigma_\epsilon^2} \\
 &\quad \text{since } \text{var}(X_S) = \sigma_\epsilon^2 / (1 - \beta_1^2), X_S = \beta_0 + \beta_1 X_{S-1} + \epsilon_S, \\
 &\quad \beta_0 \text{ is a constant, } X_{S-1} \text{ is independent of } \epsilon_i \text{ for } i \geq S, \text{ and } \text{var}(\epsilon_i) = \sigma_\epsilon^2 \text{ for } i \geq 1 \\
 &= \cancel{\beta_1^T \sigma_\epsilon^2} (3 - 2\beta_1^2) \beta_1^T \sigma_\epsilon^2 / (1 - \beta_1^2)
 \end{aligned}$$

$E(X_t)$ and $\text{var}(X_t)$ are independent of t . ~~cov(X_t, X_s)~~,
 $\text{cov}(X_S, X_{S+t})$ is independent of time S . X is
weakly stationary.

i) $\beta_1 = 1, X_0 = 0$

$$\begin{aligned}
 X_t &= \beta_0 \sum_{i=0}^{t-1} \beta_1^i + \beta_1^t X_0 + \sum_{i=0}^{t-1} \beta_1^i \epsilon_{t-i} \\
 &= \beta_0 + \sum_{i=0}^{t-1} \epsilon_{t-i}
 \end{aligned}$$

ii) $E(X_t) =$

$$\begin{aligned}
 &= E(\beta_0 t + \sum_{i=0}^{t-1} \epsilon_{t-i}) \\
 &= \beta_0 t + \sum_{i=0}^{t-1} E(\epsilon_{t-i}) \\
 &\quad \text{by linearity of expectations} \\
 &= \beta_0 t
 \end{aligned}$$

$\text{var}(X_t)$

$$\begin{aligned}
 &= \text{var}(\beta_0 t + \sum_{i=0}^{t-1} \epsilon_{t-i}) \\
 &= t \sigma_\epsilon^2 \\
 &\quad \cancel{\text{since } \{\epsilon_i\}_{i \geq 1} \text{ is iid for } i \geq 1}
 \end{aligned}$$

$E(X_t)$ and $\text{var}(X_t)$ are both dependent on t , X_t is non-stationary.

d) $X_{S+h|S}$

$$\begin{aligned}
 &= E(X_{S+h} | X_S, X_{S-1}, \dots) \\
 &= E(\beta_0 \sum_{i=0}^{h-1} \beta_1^i + \beta_1^h X_S + \sum_{i=0}^{h-1} \beta_1^i \epsilon_{S+h-i} | X_S, X_{S-1}, \dots) \\
 &= \beta_0 (1 - \beta_1^h) / (1 - \beta_1) + \beta_1^h X_S \\
 &\quad \text{by linearity of expectations, since } \beta_0 \text{ and } \beta_1 \text{ are constants, and } X_S \text{ is known within the conditional expectation, hence } E(\epsilon_i) = 0 \text{ for } i \geq 1, \text{ and } \epsilon_i \text{ is independent of } X_S, X_{S-1}, \dots \text{ for } i \geq S+1
 \end{aligned}$$

e) NBEF

$$\begin{aligned}
 &= E(X_{S+h} - X_{S+h|S})^2 \\
 &= E(\beta_0 \sum_{i=0}^{h-1} \beta_1^i + \beta_1^h X_S + \sum_{i=0}^{h-1} \beta_1^i \epsilon_{S+h-i} - \beta_0 \sum_{i=0}^{h-1} \beta_1^i \beta_1^h X_S)^2 \\
 &= E(\sum_{i=0}^{h-1} \beta_1^i \epsilon_{S+h-i})^2 \\
 &= \sum_{i=0}^{h-1} \beta_1^{2i} E(\epsilon_{S+h-i})^2 \\
 &\quad \text{by linearity of expectations, since } \cancel{\{\epsilon_i\}_{i \geq 1}} \text{ is iid}\\
 &\quad \text{hence } E(\epsilon_i \epsilon_j) = \text{cov}(\epsilon_i, \epsilon_j) = 0 \text{ and } E(\epsilon_i) = 0 \text{ for } i \geq 1 \\
 &\quad \text{hence } E(\epsilon_i \epsilon_j) = \text{cov}(\epsilon_i, \epsilon_j) \\
 &= \sum_{i=0}^{h-1} \beta_1^{2i} \sigma_\epsilon^2 \\
 &\quad \text{since } E(\epsilon_i) = 0 \text{ for } i \geq 1 \text{ hence } E(\epsilon_i)^2 = \text{var}(\epsilon_i) = \sigma_\epsilon^2 \\
 &\quad \text{for } i \geq 1 \\
 &= \sigma_\epsilon^2 (1 - \beta_1^{2h}) / (1 - \beta_1^2)
 \end{aligned}$$

when $|\beta_i| < 1$, β_i^{∞} or $\beta_i^2 < 1$, $\beta_i^{2n} \rightarrow 0$ as $n \rightarrow \infty$

$\text{MSFE} \rightarrow \sigma_z^2 / (1 - \beta_i^2) = \text{var}(x_i)$

when $\beta_i = 1$, MSFE is undefined.

a) $E(x_t)$

$$\begin{aligned} &= E(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}) \\ &= E(\varepsilon_t) + \theta_1 E(\varepsilon_{t-1}) + \dots + \theta_q E(\varepsilon_{t-q}) \\ &= 0 \end{aligned}$$

b) $\text{var}(x_t)$

$$\begin{aligned} &= \text{var}(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}) \\ &= \text{var}(\varepsilon_t) + \theta_1^2 \text{var}(\varepsilon_{t-1}) + \dots + \theta_q^2 \text{var}(\varepsilon_{t-q}) \\ &\quad \text{since } \{\varepsilon_i\} \text{ is iid hence } \text{cov}(\varepsilon_i, \varepsilon_j) = 0 \text{ for } i \neq j \\ &= \sigma_\varepsilon^2 \sum_{i=0}^q \theta_i^2, \text{ where } \theta_0 = 1 \end{aligned}$$

c) $\text{cov}(x_t, x_{t-h})$

$$\begin{aligned} &= \text{cov}(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \\ &\quad \varepsilon_{t-h} + \theta_1 \varepsilon_{t-h-1} + \dots + \theta_q \varepsilon_{t-h-q}) \\ &= 0 \end{aligned}$$

by bilinearity of covariance, since $\{\varepsilon_i\}$ is iid hence
 $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$ for $i \neq j$

d) $\text{cov}(x_t, x_{t+h})$

$$\begin{aligned} &= \text{cov}(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}, \varepsilon_{t+h} + \theta_1 \varepsilon_{t+h-1} + \theta_2 \varepsilon_{t+h-2}) \\ &= 0 \text{ if } h \geq 2 \\ &\quad \theta_2 \text{var}(\varepsilon_{t-2}) = \theta_2 \sigma_\varepsilon^2 \text{ if } h=2 \\ &\quad \theta_1 \text{var}(\varepsilon_{t-1}) + \theta_1 \theta_2 \text{var}(\varepsilon_{t-2}) = \theta_1(1+\theta_2) \sigma_\varepsilon^2 \text{ if } h=1 \\ &\quad \text{var}(\varepsilon_t) + \theta_1^2 \text{var}(\varepsilon_{t-1}) + \theta_2^2 \text{var}(\varepsilon_{t-2}) - (\theta_1^2 + \theta_2^2) \sigma_\varepsilon^2 \text{ if } h=0 \end{aligned}$$

Since $\{\varepsilon_i\}$ is iid, ε_t is independent of ε_{t+h} if $h \geq 1$, $t \neq t+h$ and $t \neq t+h-1$, hence ε_t is indep then
 ε_{t+h} and ε_{t+h-1} are independent of ε_t . Since
 x_{t+h} is a linear function of ε_{t+h} and ε_{t+h-1} ,
 x_{t+h} is independent of ε_t (and ε_t is
independent of x_{t+h})

ii) $x_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}$

$$\begin{aligned} \varepsilon_t &= x_t - \theta_1 \varepsilon_{t-1} \\ &= x_t - \theta_1(x_{t-1} - \theta_1 \varepsilon_{t-2}) \\ &= x_t - \theta_1 x_{t-1} - \theta_1^2 \varepsilon_{t-2} \end{aligned}$$

..

$$\begin{aligned} &= x_t + (-\theta_1)^1 x_{t-1} + (-\theta_1)^2 x_{t-2} + \dots + (-\theta_1)^{t-1} x_1 + (-\theta_1)^t \varepsilon_0 \\ &= x_t + (-\theta_1)^1 x_{t-1} + (-\theta_1)^2 x_{t-2} + \dots + (-\theta_1)^{t-1} x_1 \\ &= \sum_{i=0}^{t-1} (-\theta_1)^i x_{t-i} \end{aligned}$$

iii) $x_{T+1|T}$

$$\begin{aligned} &:= E(x_{T+1} | x_T, x_{T-1}, \dots) \\ &= E(\varepsilon_{T+1} + \theta_1 \varepsilon_T | x_T, x_{T-1}, \dots) \\ &= E(\varepsilon_{T+1} | x_T, x_{T-1}, \dots) + \theta_1 E(\varepsilon_T | x_T, x_{T-1}, \dots) \\ &= \theta_1 \sum_{i=0}^{T-1} (-\theta_1)^i x_{T-i} \end{aligned}$$

since $E(\varepsilon_{T+1}) = 0$ and ε_{T+1} is independent of x_T, x_{T-1}, \dots



- i) The trajectory of the time series plot is erratic and returns to the mean rapidly. Autocorrelation decays rapidly to zero as the interval increases. The time series has low persistence.



The time series appears to be non-stationary in the sample period. There appears to be stationarity in the epoch from 1945-2000. Non-stationarity in the periods prior is plausibly due to the great depression and second world war.

There does not seem to be any transformation of the data such that the estimated AR(p)-model would better forecast this time series.

- ii) The trajectory of the time series plot is relatively smooth, and takes lengthier excursions from the mean. Autocorrelation decays gradually to zero as the interval increases. The time series has high persistence.



The time series appears to be non-stationary in the sample period. There appears to be stationarity in the epochs from 1960 to 1973 and from 1987 to 2000.

An estimated AR(p)-model of ~~short period-on-period change in the dividend-price ratio~~



~~There does not seem to be any transformation of the data such that the estimated AR(p)-model could better forecast the time series.~~

- iii) The trajectory of the time series plot is relatively smooth, and takes lengthier excursions from the mean. Autocorrelation decays gradually to zero as the interval increases. The time series has high persistence.

The time series appears to be non stationary in the sample period. There do not appear to be any ~~identifi~~ epochs where the time series is stationary.

~~There does not seem to be any transformation of the data such that the estimated AR(p)-model could better forecast the time series.~~

- iv) The trajectory of the time series plot is erratic and returns to the mean rapidly. Autocorrelation decays rapidly to zero as the interval increases. The time series has low persistence.

The time series appears to be non stationary in the sample period. There appears to be

Estimating an AR(p)-model of the period-on-period change could yield a better forecast since it appears to be more plausible that the time series of period-on-period change is stationary.



Estimating an AR(p)-model of the period-on-period change could yield a better forecast since it appears to be more plausible that the time series of period-on-period changes is stationary.

statonarity in the epoch from 1960 onwards.
Non-stationarity in the prior periods is plausibly explained by the Second World War.

There does not seem to be any transformation of the data such that the estimated AR(p) model would better forecast the time series.

4a No.

$$Y'_t = 1200 \cdot (P_t - P_{t-1}) / P_{t-1}$$

gives the monthly percentage change in industrial production, measured in percentage points per annum



$$\begin{aligned} b) \hat{Y}_{2013M1|2013M12} &= 0.79 + 0.05 Y_{2013M12} + 0.19 Y_{2013M11} \\ &\quad + 0.23 Y_{2013M10} + 0.16 Y_{2013M9} \\ &= 0.79 + 0.05(101.4) + 0.19(101.0) \\ &\quad + 0.23(100.4) + 0.16(100.2) \\ &= \end{aligned}$$

$$Y_{2013M12} = 1200 \ln(101.4 / 101.0) = 4.7431$$

$$Y_{2013M11} = 1200 \ln(101.0 / 100.4) = 7.1500$$

$$Y_{2013M10} = 1200 \ln(100.4 / 100.2) = 2.3928$$

$$Y_{2013M9} = 1200 \ln(100.2 / 99.6) = 7.2072$$

$$\begin{aligned} \hat{Y}_{2014M1|2013M12} &= 0.79 + 0.05 Y_{2013M12} + 0.19 Y_{2013M11} \\ &\quad + 0.23 Y_{2013M10} + 0.16 Y_{2013M9} \\ &= 4.0892 \end{aligned}$$

$$\begin{aligned} \hat{Y}_{2014M2|2013M12} &= 0.79 + 0.05 \hat{Y}_{2013M11|2013M12} \\ &\quad + 0.19 Y_{2013M12} + 0.23 Y_{2013M11} \\ &\quad + 0.16 Y_{2013M10} \\ &= 3.9280 \end{aligned}$$

- c) The forecaster might have suspected that industrial production P_t and hence Y_t are exhibit seasonality with a one-year interval, hence that there is some relationship between Y_t and Y_{t-12} . Feasible reasons for seasonality
 - Industrial production might exhibit seasonality because, for example, industrial commodities are used as inputs for the production of consumer goods and demand for consumer goods is highly seasonal, with peaks in the holiday period.

$$H_0: \beta_{12} = 0$$

$$H_1: \beta_{12} \neq 0$$



$$p\text{-value} = 2\phi(-0.063)$$

$$\begin{aligned} p\text{-value} &= 2\phi(-|\hat{\beta}_{12}| / \text{se}(\hat{\beta}_{12})) \\ &= 2\phi(-0.063 / 0.045) \\ &= 0.1615 \end{aligned}$$

Fail to reject the null hypothesis that the population regression parameter (coefficient of Y_{t-12}) in a regression of Y_t on $Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-4}, Y_{t-12}$ is zero at all levels of significance below 0.1615 (including all conventional levels of significance). Evidence that Y_t is related to Y_{t-12} is not strong. This variable should not be included in retained in the forecasting model.

Is this language accurate?

$$d) AIC_m := \ln(SSR_m/T) + m(2/T)$$

$$BIC_m := \ln(SSR_m/T) + m(\ln T/T)$$

$$m = p+1$$

$$T = 28 \times 12 = 336$$

$$BIC_1 = \ln(19533/336) + 2(2/336) =$$

$$BIC_2 = \ln(18683/336) + 3(2/336) =$$

$$BIC_3 = \ln(17377/336) + 4(2/336) =$$

$$BIC_4 = \ln(16285/336) + 5(2/336) = 4.0801$$

$$BIC_5 = \ln(15842/336) + 6(2/336) = 4.0507$$

$$BIC_6 = \ln(15824/336) + 7(2/336) = 3.9977$$

$$BIC_7 = \ln(15824/336) + 8(2/336) = 3.9501$$

$$BIC_8 = \ln(15824/336) + 9(2/336) = 3.9399$$

$$BIC_9 = \ln(15824/336) + 10(2/336) = 3.9560$$

$$BIC_{10} = \ln(15824/336) + 11(2/336) = 3.9734$$

~~BIC~~ $m=5$ minimises BIC_m . An AR(4) model is optimal by the BIC.

$$AIC_1 = \ln(19533/336) + 2(2/336) = 4.0687$$

$$AIC_2 = \ln(18683/336) + 3(2/336) = 4.0280$$

$$AIC_3 = \ln(17377/336) + 4(2/336) = 3.9636$$

$$AIC_4 = \ln(16285/336) + 5(2/336) = 3.9047$$

$$AIC_5 = \ln(15842/336) + 6(2/336) = 3.8831$$

$$AIC_6 = \ln(15824/336) + 7(2/336) = 3.8819$$

$$AIC_7 = \ln(15824/336) + 8(2/336) = 3.8938$$

$m=5$ minimises AIC_m . An AR(4) model is optimal by the AIC. The results do not differ.

- e) One variable Granger-causes another iff the former lags of the former carry useful information about the latter in addition to that carried by the lags of the latter such that the mean squared error made by the forecast of the latter on the basis of both variables is less than that made by the forecast of the latter on the basis of the latter alone.

consider the ADX (AR(4)) model for $\{\Delta Y_t\}$

$$\begin{aligned} Y_t = & \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \beta_3 Y_{t-3} + \beta_4 Y_{t-4} \\ & + \beta_5 \Delta R_{t-1} + \beta_6 \Delta R_{t-2} + \beta_7 \Delta R_{t-3} + \beta_8 \Delta R_{t-4} \\ & + u_t \end{aligned}$$

where $E(u_t) = 0$, ~~and~~ $\text{cov}(u, Y_{t-1}) = \dots = \text{cov}(u, Y_{t-4}) = 0$,
 $\text{cov}(u, \Delta R_{t-1}) = \dots = \text{cov}(u, \Delta R_{t-4}) = 0$

$$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$

$$H_1: \text{At least one of } \beta_1, \dots, \beta_4 \neq 0$$

$$F\text{-statistic} = ((SSR_{AR(4)} - SSR_{AR(0)})/q) / (SSR_{AR(0)} / (n-k-1))$$

where $q=4$, $n=T=336$, $k=8$, $SSR_{AR(4)} = 15824$,

$$SSR_{AR(0)} = 13147$$

$$F = 16.758$$

Under the null, $F \xrightarrow{d} F_{4,327}$

- f) I_P^t and R_t^t appear to be non-stationary, while γ_t and ΔR_t^t appear to be more plausibly stationary. I_P^t and R_t^t appear to trend upwards and downwards respectively in the sample period, such that the expected value is time-dependent.

$$\text{So } IC := \ln(\text{SSR}/\tau) + M(\bar{M}/\tau)$$

Consider the AR(p,q) models

$$Y_t = \beta_0 + \sum_{i=1}^p \beta_i Y_{t-i} + \sum_{i=1}^q \beta_i X_{t-i} + u_t$$

$$\text{where } E(u_t) = 0, \text{cov}(Y_t, u) = \dots = \text{cov}(Y_p, u) = 0, \text{cov}(X_t, u) = 0.$$

$$\text{so if } \cancel{\beta} \rightarrow \text{cov}(\beta q, u) = 0$$

$$\text{where } E(u) = 0, \text{cov}(Y_{t-1}, u) = \dots = \text{cov}(Y_{t-p}, u) = 0,$$

$$\text{cov}(X_{t-1}, u) = \dots = \text{cov}(X_{t-q}, u) = 0 \text{ by construction}$$

$$\hat{\alpha}_t = \beta_0 + \sum_{i=1}^p \hat{\beta}_i \hat{Y}_{t-i} + \sum_{i=1}^q \hat{\beta}_i \hat{X}_{t-i} + \hat{u}_t$$

$$\text{where } E(\hat{u}_t) = 0, \text{cov}(\hat{\alpha}_{t-1}, \hat{u}_t) = \dots = \text{cov}(\hat{\alpha}_{t-p}, \hat{u}_t) = 0,$$

$$\text{cov}(\hat{X}_t, \hat{u}_t) = \dots = \text{cov}(\hat{b} \hat{X}_{t-1}, \hat{u}_t) = \dots = \text{cov}(\hat{b} \hat{X}_{t-p}, \hat{u}_t) = 0 \text{ by}$$

construction.

$$\hat{u}_t = Y_t - \hat{\beta}_0 - \sum_{i=1}^p \hat{\beta}_i \hat{Y}_{t-i} - \sum_{i=1}^q \hat{\beta}_i \hat{X}_{t-i}$$

$$\hat{\beta}'_i = \frac{\text{cov}(\hat{\alpha}_t, \hat{\alpha}_{t-i})}{\text{var}(\hat{\alpha}_{t-i})}$$

For $1 \leq i \leq p$

$$\hat{\beta}'_i = \frac{\text{cov}(\hat{\alpha}_t, \hat{\alpha}_{t-i})}{\text{var}(\hat{\alpha}_{t-i})}$$

$$= \frac{\alpha^2 \text{cov}(Y_t, \hat{Y}_{t-i})}{\text{var}(\hat{Y}_{t-i})}$$

to

where \hat{Y}_{t-i} is the residual in the regression of

Y_{t-i} on $Y_{t-1}, \dots, Y_{t-i+1}, Y_{t-i-1}, \cancel{\dots}, Y_{t-p}, X_{t-1},$

\dots, X_{t-q}

$$= \frac{\alpha^2 \text{cov}(Y_t, \hat{Y}_{t-i})}{\alpha^2 \text{var}(\hat{Y}_{t-i})}$$

$$= \frac{\text{cov}(Y_t, \hat{Y}_{t-i})}{\text{var}(\hat{Y}_{t-i})}$$

$$= \hat{\beta}_i$$

For $1 \leq i \leq q$

$$\hat{\beta}'_i = \frac{\text{cov}(\hat{\alpha}_t, b \hat{X}_{t-i})}{\text{var}(b \hat{X}_{t-i})}$$

where \hat{X}_{t-i} is the residual in the regression of

X_{t-i} on $Y_{t-1}, \dots, Y_{t-p}, X_{t-1}, \dots, X_{t-i+1}, X_{t-i-1}, \dots,$

X_{t-q}

$$= \frac{abc \text{cov}(Y_t, \hat{X}_{t-i})}{b^2 \text{var}(\hat{X}_{t-i})}$$

$$= (\hat{\alpha}_0) \hat{\beta}_i$$

$$\cancel{\hat{\beta}_0} = \cancel{\hat{\beta}_0}$$

$$\hat{\beta}_0 = \bar{Y} - \sum_{i=1}^p \hat{\beta}_i \bar{Y}_{t-i} - \sum_{i=1}^q \hat{\beta}_i \bar{X}_{t-i}$$

$$\hat{\beta}_0 = \bar{Y} - \sum_{i=1}^p \hat{\beta}_i \bar{Y}_{t-i} - \sum_{i=1}^q \hat{\beta}_i b \bar{X}_{t-i}$$

$$= \bar{Y} - a \sum_{i=1}^p \hat{\beta}_i \bar{Y}_{t-i} - a \sum_{i=1}^q \hat{\beta}_i \bar{X}_{t-i}$$

$$= a \hat{\beta}_0$$

$$\hat{u}'_t = \hat{\alpha}_t - \hat{\beta}_0' - \sum_{i=1}^p \hat{\beta}_i' \hat{Y}_{t-i} - \sum_{i=1}^q \hat{\beta}_i' b \hat{X}_{t-i}$$

$$= \hat{\alpha}_t - a \hat{\beta}_0 - a \sum_{i=1}^p \hat{\beta}_i \hat{Y}_{t-i} - a \sum_{i=1}^q \hat{\beta}_i b \hat{X}_{t-i}$$

$$= a \hat{u}_t$$

$$\text{SSR}' = \sum_{i=1}^T (\hat{u}'_t)^2 = \sum_{i=1}^T (a \hat{u}_t)^2 = a^2 \sum_{i=1}^T \hat{u}_t^2 = a^2 \text{SSR}$$

The decrease in the term $\ln(\text{SSR}/\tau)$ associated with an increase in M is greater for $a > 1$ in the re-estimated model if $|a| > 1$ hence $a^2 > 1$ and smaller if $|a| < 1$ hence $a^2 < 1$. Assuming that the re-estimation has no effect on \bar{M} , the increase in the term $M(\bar{M}/\tau)$ associated with

An increase in m is unchanged from the original model to the reestimated model. Therefore, if $|a| > 1$, the decrease in the former term is likely to be greater than the increase in the latter term, ~~over a greater up to higher values of m in the re-estimated model as opposed to the original, hence minimisation of ℓ_k yields a re-estimated model is likely to yield a re-estimated model containing more terms of T and X , than the original model. The reverse is true if $|a| < 1$.~~