QE Problem Set 7

1a $x_{s+t} = \beta_0 + \beta_1 x_{s+t-1} + \varepsilon_{s+t}$

$= \beta_0 + \beta_1(\beta_0 + \beta_1 x_{s+t-2} + \varepsilon_{s+t-1}) + \varepsilon_{s+t}$

$= (\beta_0 + \beta_1\beta_0) + \beta_1^2 x_{s+t-2} + \varepsilon_{s+t} + \beta_1 \varepsilon_{s+t-1}$

$= \dots$

$= \cancel{\beta_0(1+\beta_1+\beta_1^2+\dots)+\beta_1}$

$\beta_0(1+\beta_1+\beta_1^2+\dots+\beta_1^{t-1})$

$+\beta_1^t x_s$

$+(\varepsilon_{s+t} + \beta_1\varepsilon_{s+t-1} + \beta_1^2\varepsilon_{s+t-2} + \dots + \beta_1^{t-1}\varepsilon_{s+1})$

$= \beta_0 \sum_{i=0}^{t-1} \beta_1^i + \beta_1^t x_s + \sum_{i=0}^{t-1} \beta_1^i \varepsilon_{s+t-i}$

$s = 0$

$x_t = \beta_0 \sum_{i=0}^{t-1} \beta_1^i + \beta_1^t x_0 + \sum_{i=0}^{t-1} \beta_1^i \varepsilon_{t-i}$

~~Given that~~ Since β_0 and β_1 are constants, $\beta_0 \sum_{i=0}^{t-1} \beta_1^i$ is independent of ε_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, $\beta_1^t x_0$ is independent of ε_i for all $i \geq 1$.
Dependence of x_t on $\cancel{\varepsilon}$ $\{\varepsilon_i\}$ is given entirely by the term $\sum_{i=0}^{t-1} \beta_1^i \varepsilon_{t-i}$. x_t depends directly only on $\{\varepsilon_t, \varepsilon_{t-1}, \dots \varepsilon_1\}$. Since $\{\varepsilon_i\}$ is iid, x_t is independent of $\{\varepsilon_{t+1}, \varepsilon_{t+2}, \dots\}$.

b $E(x_t)$

$= E(\beta_0 \sum_{i=0}^{t-1} \beta_1^i + \beta_1^t x_0 + \sum_{i=0}^{t-1} \beta_1^i \varepsilon_{t-i})$

$= \beta_0(1-\beta_1^t)/(1-\beta_1) + \beta_1^t E(x_0) + \sum_{i=0}^{t-1} \beta_1^i E(\varepsilon_{t-i})$

by linearity of expectations, since β_0 and β_1 are constants

$= \beta_0(1-\beta_1^t)/(1-\beta_1) + \beta_1^t \beta_0/(1-\beta_1)$

given $E(x_0) = \beta_0/(1-\beta_1)$ and $E(\varepsilon_i) = 0$ for $i \geq 1$

$= \beta_0/(1-\beta_1)$

$var(x_t)$

$= var(\beta_0 \sum_{i=0}^{t-1} \beta_1^i + \beta_1^t x_0 + \sum_{i=0}^{t-1} \beta_1^i \varepsilon_{t-i})$

$= \beta_1^{2t} var(x_0) + \sum_{i=0}^{t-1} \beta_1^i var(\varepsilon_{t-i})$

since β_0 and β_1 are constants, x_0 is independent of ε_i for $i \geq 1$, and $\cancel{\varepsilon}$ $\{\varepsilon_i\}$ is iid

$= \beta_1^{2t} \sigma_\varepsilon^2/(1-\beta_1^2) + \sum_{i=0}^{t-1} \beta_1^{2i} \sigma_\varepsilon^2$

given $var(x_0) = \sigma_\varepsilon^2/(1-\beta_1^2)$ and $var(\varepsilon_i) = \sigma_\varepsilon^2$ for all $i \geq 1$

$= \beta_1^{2t}\sigma_\varepsilon^2(1-\beta_1^2) + (1-\beta_1^{2t})\sigma_\varepsilon^2/(1-\beta_1^2)$

$= \sigma_\varepsilon^2/(1-\beta_1^2)$

$cov(x_s, x_{s+t})$

$= cov(\beta_0 \sum_{i=0}^{-1} \beta_1^i + \beta_1^0 x_s + \sum_{i=0}^{-1} \beta_1^i \varepsilon_{s-i},$

$\beta_0 \sum_{i=0}^{t-1} \beta_1^i + \beta_1^t x_s + \sum_{i=0}^{t-1} \beta_1^i \varepsilon_{s+t-i})$

$= cov(\cancel{\beta} x_s + \sum_{i=0}^{-1} \beta_1^i \varepsilon_{s-i}, \beta_1^t x_s + \sum_{i=0}^{t-1} \beta_1^i \varepsilon_{s+t-i})$

$\cancel{= \beta_1^t var(x_s) + cov(\varepsilon_s + \beta_1^{-1}\varepsilon_{s+1}, \beta_1^0 \varepsilon_{s+t}}$

$\cancel{+ cov(\varepsilon_s + \beta_1^{-1}\varepsilon_{s+1}, \beta_1^0 \varepsilon_{s+t} + \dots + \beta_1^{t-1}\varepsilon_{s+1})}$

~~since x_s is independent of ε_i for $i \geq s+1$~~

since β_0 and β_1 are constants

$= cov(x_s + \varepsilon_s + \beta_1^{-1}\varepsilon_{s+1}, \beta_1^t x_s + \sum_{i=0}^{t-1} \beta_1^i \varepsilon_{s+t-i})$

$\cancel{= cov(x_s + \varepsilon_s, \beta_1^t x_s + \varepsilon_{s+t} + \dots + \beta_1^{t-1}\varepsilon_{s+1})}$

$= cov(X_s + \varepsilon_s + \beta_1^{-1}\varepsilon_{s+1}, \beta_1^t X_s + \varepsilon_{s+t} + \cdots + \beta_1^{t+1}\varepsilon_{s+1})$

$= \beta_1^t var(X_s) + \beta_1^t cov(X_s, \varepsilon_s) + \beta_1^t var(\varepsilon_{s+1})$

by bilinearity of covariance, since X_s is independent of ε_i for $i \geq s+1$ and $\{\varepsilon_i\}$ is iid.

$= \beta_1^t \sigma_\varepsilon^2/(1-\beta_1^2)$ ~~$+2\beta_1^t var$~~ $2\beta_1^t \sigma_\varepsilon^2$

since $var(X_s) = \sigma_\varepsilon^2/(1-\beta_1^2)$, $X_s = \beta_0 + \beta_1 X_{s-1} + \varepsilon_s$, β_0 is a constant, X_{s-1} is independent of ε_i for $i \geq s$, and $var(\varepsilon_i) = \sigma_\varepsilon^2$ for $i \geq 1$

$=$ ~~$(3-2\beta_1^2$~~ $(3-2\beta_1^2)\beta_1^t \sigma_\varepsilon^2/(1-\beta_1^2)$

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

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

$X_t = \beta_0 \sum_{i=0}^{t-1} \beta_1^i + \beta_1^t X_0 + \sum_{i=0}^{t-1} \beta_1^i \varepsilon_{t-i}$

$= \beta_0 t + \sum_{i=0}^{t-1} \varepsilon_{t-i}$

ii $E(X_t)$

$= E(\beta_0 t + \sum_{i=0}^{t-1} \varepsilon_{t-i})$

$= \beta_0 t + \sum_{i=0}^{t-1} E(\varepsilon_{t-i})$

by linearity of expectations

$= \beta_0 t$

$var(X_t)$

$= var(\beta_0 t + \sum_{i=0}^{t-1} \varepsilon_{t-i})$

$= t\sigma_\varepsilon^2$

~~by~~ since $\{\varepsilon_i\}$ is iid for $i \geq 1$

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

d $X_{s+h|s}$

$:= E(X_{s+h} \mid X_s, X_{s-1}, \ldots)$

$= E(\beta_0 \sum_{i=0}^{h-1} \beta_1^i + \beta_1^h X_s + \sum_{i=0}^{h-1} \beta_1^i \varepsilon_{s+h-i} \mid X_s, X_{s-1}, \ldots)$

$= \beta_0(1-\beta_1^h)/(1-\beta_1) + \beta_1^h X_s$

by linearity of expectations, since β_0 and β_1 are constants, ~~and~~ X_s is known within the conditional expectation, ~~and~~ $E(\varepsilon_i) = 0$ for $i \geq 1$, and ε_i is independent of $X_s, X_{s-1}, \ldots$ for $i \geq s+1$

e MSFE

$:= E(X_{s+h} - X_{s+h|s})^2$

$= E(\beta_0 \sum_{i=0}^{h-1} \beta_1^i + \beta_1^h X_s + \sum_{i=0}^{h-1} \beta_1^i \varepsilon_{s+h-i} - \beta_0 \sum_{i=0}^{h-1} \beta_1^i + \beta_1^h X_s)^2$

$= E(\sum_{i=0}^{h-1} \beta_1^i \varepsilon_{s+h-i})^2$

$= \sum_{i=0}^{h-1} \beta_1^{2i} E(\varepsilon_{s+h-i})^2$

by linearity of expectations, since ~~given~~ $\{\varepsilon_i\}$ is iid ~~hence $E(\varepsilon_i\varepsilon_j) = cov(\varepsilon_i, \varepsilon_j) = 0$~~ and $E(\varepsilon_i) = 0$ for $i \geq 1$ hence $E(\varepsilon_i \varepsilon_j) = cov(\varepsilon_i, \varepsilon_j)$

$= \sum_{i=0}^{h-1} \beta_1^{2i} \sigma_\varepsilon^2$

since $E(\varepsilon_i) = 0$ for $i \geq 1$ ~~he~~ hence $E(\varepsilon_i)^2 = var(\varepsilon_i) = \sigma_\varepsilon^2$ for $i \geq 1$

$= \sigma_\varepsilon^2(1-\beta_1^{2h})/(1-\beta_1^2)$

when $|\beta_1|<1$, $\cancel{\beta_1^2}$ as $\beta_1^2<1$, $\beta_1^{2h} \to 0$ as $h \to \infty$
MSFE $\to \sigma_z^2/(1-\beta_1^2)$ ~~= var(Yt)~~

when $\beta_1 = 1$, MSFE is undefined.

2a $E(X_t)$
$= E(\varepsilon_t + \theta_1\varepsilon_{t-1} + \dots + \theta_q\varepsilon_{t-q})$
$= E(\varepsilon_t) + \theta_1 E(\varepsilon_{t-1}) + \dots + \theta_q E(\varepsilon_{t-q})$
$= 0$

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

b $cov(X_t, X_{t-h})$
$= cov(\varepsilon_t + \theta_1\varepsilon_{t-1} + \dots + \theta_q\varepsilon_{t-q},$
$\varepsilon_{t-h} + \theta_1\varepsilon_{t-h-1} + \dots + \theta_q\varepsilon_{t-h-q})$
$= 0$
by bilinearity of covariance, since $\{\varepsilon_i\}$ is iid hence $cov(\varepsilon_i, \varepsilon_j) = 0$ for $i \neq j$

c $cov(X_t, X_{t-h})$
$= cov(\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2}, \varepsilon_{t-h} + \theta_1\varepsilon_{t-h-1} + \theta_2\varepsilon_{t-h-2})$
$= 0$ if $h > 2$
$\theta_2 var(\varepsilon_{t-2}) = \theta_2\sigma_\varepsilon^2$ if $h = 2$
$\theta_1 var(\varepsilon_{t-1}) + \theta_1\theta_2 var(\varepsilon_{t-2}) = \theta_1(1+\theta_2)\sigma_\varepsilon^2$ if $h = 1$
$var(\varepsilon_t) + \theta_1^2 var(\varepsilon_{t-1}) + \theta_2^2 var(\varepsilon_{t-2}) = (1+\theta_1^2+\theta_2^2)\sigma_\varepsilon^2$ if $h = 0$

d i since $\{\varepsilon_t\}$ is iid, ~~ε_t is independent of~~ and ~~for~~ 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, ~~then~~ (and ε_t is independent of X_{t-h})

ii $X_t = \varepsilon_t + \theta_1\varepsilon_{t-1}$
$\varepsilon_t = X_t - \theta_1\varepsilon_{t-1}$
$= X_t - \theta_1(X_{t-1} - \theta_1\varepsilon_{t-2})$
$= X_t - \theta_1 X_{t-1} - \theta_1^2\varepsilon_{t-2}$
$= \dots$
$= X_t + (-\theta_1)^1 X_{t-1} + (-\theta_1)^2 X_{t-2} + \dots + (-\theta_1)^{t-1} X_1 + (-\theta_1)^t \varepsilon_0$
$= X_t + (-\theta_1)^1 X_{t-1} + (-\theta_1)^2 X_{t-2} + \dots + (-\theta_1)^{t-1} X_1$
$= \sum_{i=0}^{t-1} (-\theta_1)^i X_{t-i}$

iii $X_{T+1|T}$
$:= E(X_{T+1} | X_T, X_{T-1}, \dots)$
$= E(\varepsilon_{T+1} + \theta_1\varepsilon_T | X_T, X_{T-1}, \dots)$
$= E(\varepsilon_{T+1} | X_T, X_{T-1}, \dots) + \theta_1 E(\varepsilon_T | X_T, X_{T-1}, \dots)$
$= \theta_1 \sum_{i=0}^{t-1} (-\theta_1)^i X_{t-i}$
since $E(\varepsilon_{T+1}) = 0$ and ε_{T+1} is independent of X_T, $X_{T-1}, \dots$

3i 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 could 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 ~~each~~ 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.~~

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

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.~~

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

iv The trajectory of the time series plot is erratic and returns to the mean rapidly. Autocorrelation decays rapidly to zero as the interval increase. 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 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.

$Y'_t = 1200\ (IP_t - IP_{t-1}) / IP_{t-1}$

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

b ~~$\hat{Y}_{2014M1|2013M12} = 0.79 + 0.05Y_{2013M12} + 0.19Y_{2013M11} + 0.23Y_{2013M10} + 0.16Y_{2013M9}$~~
~~$= 0.79 + 0.05(101.4) + 0.19(101.0) + 0.23(100.4) + 0.16(100.2)$~~

$Y_{2013m12} = 1200 \ln(101.4/101.0) = 4.7431$
$Y_{2013M11} = 1200 \ln(101.0/100.4) = 7.1500$
$Y_{2013M10} = 1200 \ln(100.4/100.2) = 2.3928$
$Y_{2013M9} = 1200 \ln(100.2/99.6) = 7.2072$

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

$\hat{Y}_{2014M2|2013M12} = 0.79 + 0.05\hat{Y}_{2014M1|2013M12} + 0.19Y_{2013M12} + 0.23Y_{2013M11} + 0.16Y_{2013M10}$
$= 3.9280$

c The forecaster might have suspected that industrial production IP and hence Y ~~are~~ exhibit seasonality with a one-year interval, hence that there is some relationship between Y_t and Y_{t-12}. ~~Plausible reasons for seasonality include~~ 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.

$H_0: \beta_{12} = 0$
$H_1: \beta_{12} \neq 0$

~~$\text{p-value} = 2\Phi^{-1}(-0.063)$~~
$\text{p-value} = 2\Phi^{*}(-|\hat{\beta}_{12}| / \hat{se}(\hat{\beta}_{12}))$
$= 2\Phi^{*}(-0.063/0.045)$
$= 0.16151$

Fail to reject the null hypothesis that the ~~populatio~~ 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 0.16151 (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 $AIC_m := \ln(SSR_m/T) + m(2/T)$
$BIC_m := \ln(SSR_m/T) + m(\ln T/T)$
$m = p+1$
$T = 28 \times 12 = 336$

~~$BIC_1 = \ln(19533/336) + m(2/336) =$~~
~~$BIC_2 = \ln(18643/336) + 2(2/336) =$~~
~~$BIC_3 = \ln(17377/336) +$~~

$BIC_1 = \ln(19533/336) + \ln 336/336 = 4.0801$
$BIC_2 = \ln(18643/336) + 2(\ln 336/336) = 4.0507$
$BIC_3 = \ln(17377/336) + 3(\ln 336/336) = 3.9977$
$BIC_4 = \ln(16285/336) + 4(\ln 336/336) = 3.9501$
$BIC_5 = \ln(15842/336) + 5(\ln 336/336) = 3.9399$
$BIC_6 = \ln(15824/336) + 6(\ln 336/336) = 3.9560$
$BIC_7 = \ln(15824/336) + 7(\ln 336/336) = 3.9734$

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

$AIC_1 = \ln(19533/336) + 2/336 = 4.0687$
$AIC_2 = \ln(18643/336) + 2(2/336) = 4.0280$
$AIC_3 = \ln(17377/336) + 3(2/336) = 3.9636$
$AIC_4 = \ln(16285/336) + 4(2/336) = 3.9047$
$AIC_5 = \ln(15842/336) + 5(2/336) = 3.8831$
$AIC_6 = \ln(15824/336) + 6(2/336) = 3.8879$
$AIC_7 = \ln(15824/336) + 7(2/336) = 3.8938$

$m=5$ 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 ADL(4,4) model for $\{Y_t\}$
$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \beta_3 Y_{t-3} + \beta_4 Y_{t-4} + \gamma_1 \Delta R_{t-1} + \gamma_2 \Delta R_{t-2} + \gamma_3 \Delta R_{t-3} + \gamma_4 \Delta R_{t-4} + u_t$$
where $E(u_t)=0$, $cov(u, Y_{t-1}) = \dots = cov(u, Y_{t-4}) = 0$, $cov(u, \Delta R_{t-1}) = \dots = cov(u, \Delta R_{t-4}) = 0$

$H_0: \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0$
H_1: At least one of $\gamma_1, \dots, \gamma_4 \neq 0$

F-statistic: $((SSR_{rs} - SSR_{un})/q)/(SSR_{un}/(n-k-1))$
where $q=4$, $n=T=336$, $k=8$, $SSR_{rs} = 15842$, $SSR_{un} = 13147$
$F = 16.758$
under the null, $F \xrightarrow{d} F_{4,\infty}$

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

5a $IC := \ln(SSR_m/T) + m(\eta_T/T)$

Consider the ADL(p,q) models

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

~~where $E(u_t)=0$, $cov(Y_1,u)=\dots=cov(Y_p,u)=0$, $cov(X_1,u)=$ $cov(X\dots = cov(X_q,u)=0$~~

where $E(u)=0$, $cov(Y_{t-1},u)=\dots=cov(Y_{t-p},u)=0$, $cov(X_{t-1},u)=\dots=cov(X_{t-q},u)=0$ by construction

$aY_t = \beta_0' + \sum_{i=1}^{p} \beta_i' aY_{t-i} + \sum_{i=1}^{q} \delta_i' bX_{t-i} + u_t'$

where $E(u')=0$, $cov(aY_{t-1},u')=\dots=cov(aY_{t-p},u')=0$, ~~$cov(X_{t-}$~~ $cov(bX_{t-1},u')=\dots=cov(bX_{t-p},u')=0$ by construction.

$\hat{u}_t = Y_t - \hat{\beta}_0 - \sum_{i=1}^{p} \hat{\beta}_i Y_{t-i} - \sum_{i=1}^{q} \hat{\delta}_i X_{t-i}$

~~$\hat{\beta}_i' = \hat{cov}(aY_t, aY_{t-i})/\hat{var}$~~

For $1 \leq i \leq p$

$\hat{\beta}_i' = \hat{cov}(aY_t, a\tilde{Y}_{t-i}) / \hat{var}(a\tilde{Y}_{t-i})$

~~$= a^2 \hat{cov}(Y_t, \tilde{Y}_{t-i})/\hat{var}($ $a^2 \hat{var}(\tilde{Y}_{t-i})$~~

~~to~~

where $\tilde{Y}_{t-i}$ is the residual in the regression of Y_{t-i} on $Y_{t-1}, \dots, Y_{t-i+1}, Y_{t-i-1}, \dots, Y_{t-p}, X_{t-1}, \dots, X_{t-q}$

~~$= a^2 \hat{cov}(aY_t, \tilde{Y}_{t-i}) / \hat{var}($~~

$= a^2 \hat{cov}(Y_t, \tilde{Y}_{t-i}) / a^2 \hat{var}(\tilde{Y}_{t-i})$

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

$= \hat{\beta}_i$

For $1 \leq i \leq q$

$\hat{\delta}_i' = \hat{cov}(aY_t, b\tilde{X}_{t-i}) / \hat{var}(b\tilde{X}_{t-i})$

where $\tilde{X}_{t-i}$ is the residual in the regression of X_{t-i} on $Y_{t-1}, \dots, Y_{t-p}, X_{t-1}, \dots, X_{t-i+1}, X_{t-i-1}, \dots, X_{t-q}$

$= ab\hat{cov}(Y_t, \tilde{X}_{t-i}) / b^2 \hat{var}(\tilde{X}_{t-i})$

$= (a/b)\hat{\delta}_i$

~~$\hat{u}_t' = aY_t$~~

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

$\hat{\beta}_0' = a\bar{Y} - \sum_{i=1}^{p} \hat{\beta}_i' a\bar{Y}_{t-i} - \sum_{i=1}^{q} \hat{\delta}_i' b\bar{X}_{t-i}$

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

$= a\hat{\beta}_0$

$\hat{u}_t' = aY_t - \hat{\beta}_0' - \sum_{i=1}^{p} \hat{\beta}_i' aY_{t-i} - \sum_{i=1}^{q} \hat{\delta}_i' bX_{t-i}$

$= aY_t - a\hat{\beta}_0 - a\sum_{i=1}^{p} \hat{\beta}_i Y_{t-i} - a\sum_{i=1}^{q} \hat{\delta}_i X_{t-i}$

$= a\hat{u}_t$

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

The decrease in the term $\ln(SSR_m/T)$ associated with an increase in m is greater ~~smaller~~ in the re-estimated model if ~~$a^2 \geq 1$~~ $|a|>1$ hence $a^2>1$ and smaller if $|a|<1$ hence $a^2<1$. Assuming that the re-estimation has no effect on η_T, the increase in the term $m(\eta_T/T)$ 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 more 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 IC ~~yields a re-estimated model~~ is likely to yield a re-estimated model containing more lags of Y and X, than the original model. The reverse is true if $|a|<1$.