

Instructions

The exam consists of Question 1 and Question 2, each one presenting a number of subquestions. On page 11 and page 13 you will find the Stata commands (do-file) and output (log-file) relative to Question 1. On page 22 and page 24 you will find the Stata commands (do-file) and output (log-file) relative to Question 2. Read carefully the text. Answer all questions. Good luck!

Question 1 (45 points)

This question uses material from Clarke and Batina (2019). “A Replication of ‘Is Public Expenditure Productive?’ (Journal of Monetary Economics, 1989).” Public Finance Review 47.3, 623-629.

Clarke and Batina (2019) replicate the analysis in Aschauer (1989). The objective of Aschauer (1989) was to empirically investigate the effects of government spending on private investment, answering the question: does higher public capital accumulation ‘crowd out’ private investment in plant and equipment? “On neoclassical grounds, the answer to this question is seen to depend upon two fundamental, opposing forces. On the one hand, higher public investment raises the national rate of capital accumulation above the level chosen (in a presumed rational fashion) by private sector agents; thus, public capital spending may crowd out private expenditures on capital goods on an ex ante basis as individuals seek to reestablish an optimal intertemporal allocation of resources. On the other hand, public capital - particularly infrastructure capital such as highways, water systems, sewers, and airports - is likely to bear a complementary relationship with private capital in the private production technology. Thus, higher public investment may raise the marginal productivity of private capital and, thereby, ‘crowd in’ private investment. ” (Aschauer, 1989. ”Is public expenditure productive?.” Journal of Monetary Economics 23.2, p. 177).

Aschauer (1989) formulates an aggregate Cobb-Douglas production function:

$$Y_t = A_t N_t^{e_N} K_t^{e_K} G_t^{e_G} \exp(\epsilon_t) \quad (1)$$

where Y_t is private output, $A_t = \exp(a_0 + a_1 t)$ is a Hicks-neutral measure of productivity, N_t is aggregate labor, K_t is aggregate private capital, G_t is public capital, ϵ_t is an error term and $(a_1, a_2, e_N, e_K, e_G)$ are the parameters to be estimated (“e” because they are elasticities). Aschauer estimates the following two equations:

$$y_t - k_t = a_0 + a_1 t + a_2(n_t - k_t) + a_3(g_t - k_t) + a_4 c u_t + \epsilon_t, \quad (2)$$

where lower case variables are in logs. The variables are:

y_t : real aggregate output (in logs)

t : a time trend

k_t : aggregate stock of capital (in logs)

t : time trend

n_t : aggregate employment (in logs)

g_t : public capital stock, e.g. stock of nonmilitary public structures and equipment (in logs)

cu_t : capacity utilization (in logs), a metric which is used to compute the rate at which probable output levels are being met or used. This variable aims to control for the influence of the business cycle.

ϵ_t : white noise disturbances, with mean zero and variance σ^2 .

The critical parameter in these equations is $a_3 (= e_G$ in equation 1), which capture the productivity of public capital.

Note that the context for this question is the same as the second in-class test. Students should therefore be familiar at least with the "story" behind the question (data were provided as well, and the suggested solutions for the test hinted to the importance of cointegration in that context although did not go into it). In addition, part (a), (b), (c), (e) are standard questions asked in class and in past exams (DF test, Johansen, trace, LM-test). Part (e) and (f) require reading of the results – but results do not present major challenges.

- (a) Explain (once) the test performed on lines 32-54 of the do-file, and provide its conclusion for all the variables (note that the critical values are reported in the do-file).

Solution. This is a basic Dickey-Fuller test. The student should report the null/alternative hypothesis, test-statistic and discuss the fact that it has a non-standard distribution. At 5% significance level, we fail to reject H_0 and therefore cannot exclude the presence of a unit root for all variables except cu_t . This excludes a cointegrating relationship between this variable and any of the others.

- (b) Define the Johansen methodology and list its purpose and advantages in the context of this application over the simple OLS estimation of line 22 of the do-file.

Solution. Students should define \mathbf{x}_t as the vector including the key variables, i.e. $y_t - k_t$, $(n_t - k_t)$, $g_t - k_t$, cu_t . Students could first define how a model such as $\mathbf{x}_t = A_0 + A_1\mathbf{x}_{t-1} + \mathbf{u}_t$ is agnostic on all cross-equation dependencies. It therefore does not exclude, as the OLS does, reverse causality and rather allows all variables to enter in all equations, as in all VARs. In addition, it allows to directly study the long-term behavior of the system by checking the existence of any cointegrating relationship between variables. Johansen's method pins down to studying the system: $\Delta\mathbf{x}_t = \boldsymbol{\pi}\mathbf{x}_t + \mathbf{u}_t$, and the properties of $\boldsymbol{\pi}$. See lecture 20, slides 6 onward. A good answer should include details on why the rank of this matrix is informative on cointegration. The perfect answer should at this point or in the next one discuss the fact that we can already expect the n. of cointegrating vectors to be less than 4 as one of the variables is stationary – as seen in part (a).

- (c) Define the trace test and comment on the results reported in the log-file (line 41).

Solution. Slide 8, lecture 20. There is one cointegrating vector.

- (d) Focus only on the key variable $(y_t - k_t)$. Interpret the results relative to this variable from the cointegration analysis reported on line 43 of the log-file.

Solution. The answer has two parts. Students should discuss both the cointegrating vector and the speed of adjustment parameters. The cointegrating relationship is:

$$\widehat{(y_t - k_t)} = -0.12(n_t - k_t) + 0.19(g_t - k_t) + 0.46cu_t + 2.14$$

Other things being equal, this relationship suggests that in equilibrium output per unit of capital (e.g. private capital productivity) is expected to decrease by 0.12% for each 1% increase in the private labor to capital ratio and increase by 0.19% for each 1% increase in the public capital to private capital ratio. This suggests that in equilibrium there is a crowding-in effect taking place. Any shock that will result in disequilibrium will cause an adjustment in the private capital productivity. This is seen in the coefficient -0.36. A 1% increase private capital productivity above its long-term value will result in a 0.36% decrease next period. The other coefficients can be interpreted as you would in a standard log-log model. It is interesting to note that public capital productivity in the short run also seem positively related to private capital productivity (0.60) – but this number is not statistically zero.

- (e) Although we have not discussed it in a VAR setting, the command `vec1mar` checks for ARCH effects, as we learned in class in the context of univariate models. Referring to a univariate

model, explain what we mean with ARCH effects and present a test that can be performed to detect them.

Solution. Students should define an ARCH model and the LM-test (a Q-test would do as well, but the LM-test is a more direct test introduced in class to discuss ARCH effects). Both tests are standard and have been asked many times, see book and notes.

- (f) Abdih and Joutz (2008) find that the residuals of a VECM have better properties when they include a dummy variable representing recessions that are usually related to energy shocks. Intuitively explain an additional, even more important reason, as to why a "recession dummy" might be important to be included in an analysis like the one presented.

Solution. We have seen in class how structural breaks can affect the results of the DF test. This is discussed on p. 227 onward of the textbook and can be seen in the test provided in the do-file: although results on the stationarity of the labor to capital ratio are not changed, the statistic is affected even just by the inclusion of a simple recession dummy. Concerns of structural breaks cause caution then in the interpretation of the results above.

- (g) Provide a critical summary of the conclusions from this analysis: is public expenditure productive?

Solution. VECM is preferred to OLS, nonetheless all models suggests a positive relationship between private and public capital. This can be seen from the OLS results (elasticity of 0.37) as well as the cointegration results (elasticity of 0.19). Nonetheless, the last point also raises concerns about possible breaks affecting the analysis. Finally, the analysis performed has not paid attention to specification issues: how were lag-length selected? All in all, results seem robust but the analysis presents room for improvement.

Question 2 (55 points)

This question is partly based on Benton, A., and A. Philips. "Does the @realDonaldTrump really matter to financial markets?", forthcoming *American Journal of Political Science* (2019).

Anecdotal evidence suggests that economic policy statements made by US President Donald J. Trump via the microblogging website Twitter have the power to rattle financial market. The academic literature however suggests that his tweets should not matter to investors. Financial economists argue that financial markets are efficient, with asset prices reflecting all publicly available information (Fama 1970, 1991). Political economists build on this to suggest that only new and unanticipated information about the future political and economic policy direction of government should affect investors' views about the future value of their assets. By this logic, Trump's economic policy tweets should have only mattered to financial markets while his policy agenda was unknown; once the direction of his economic policy views were clear, his tweets would not have provided any new information to investors, leaving financial markets untouched.

This paper examines the impact of Trump's Mexico-related policy tweets on the US dollar-Mexican peso (USD/MXN) exchange rate. This data is ideal for three reasons. First, Trump was a newcomer to national US politics, raising chances that his economic policy statements contained new information about his economic policy views. Second, during the period under examination (January 2015 to February 2018), Trump restated what are clearly negative views on the US-Mexico relationship. Third, Trump frequently expressed his Mexico-related policy views via Twitter. [...] If investors only respond to news about the likely future economic policy direction of government, then Trump's Mexico-related policy tweets should have affected the USD/MXN exchange rate early in his campaign, before his agenda became clear. If on the contrary investors respond both to news about future economic policy directions and about the future policy resolve of government, then Trump's tweets should have affected the exchange rate both before and after his Mexico-related policy views were known.

The main dependent variable is the percentage change in the daily USD/MXN from 1 January 2015 to 2 February 2018. Unit root tests indicate that the dependent variable is stationary.

Figure 1 plots both the raw dollar-peso exchange rate (e.g. 1\$ bought 14.75 pesos on Jan 1, 2015) and the % change in the dollar-peso exchange rate. The exchange rate faced periods of substantial volatility, with the most notable spike just after the 2016 US presidential election. The main explanatory variable is the daily presence of a Mexico-related policy tweet. The authors use information from the Trump Twitter Archive, which archives all tweets sent by @realDonaldTrump on GitHub. Trump sent over 14,500 tweets between 1 January 2015 and 2 February 2018.

The authors use generalized autoregressive conditional heteroskedasticity (GARCH) models for their analysis.

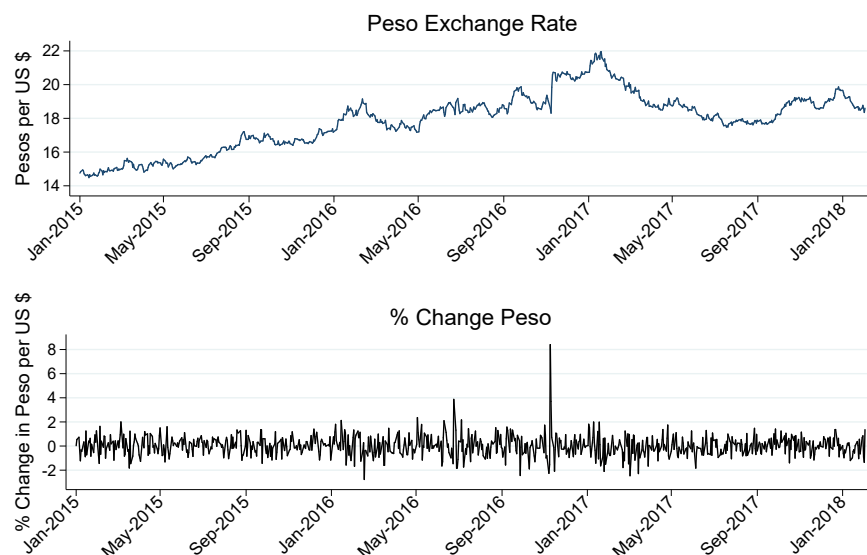


Figure 1: Peso-Dollar exchange rate

In the analysis, the variables are defined as:

`pctchange_peso`: daily % change in the nominal dollar-peso exchange rate

`tweetdum`: indicator variable taking value of 1 if Trump sent a Mexico-related tweet that day, zero otherwise

`pctchangesp500`: percentage change in the US S&P 500 stock market index, capturing shifts in expectations about US economic performance which affect views about the Mexican economy and the exchange rate

`bondspread10yr_pc`: percentage change in the 10 year Mexico-US bond spread

`lnusdstock`: is the change in the log of the Mexican central bank's reported weekly US dollar reserves.

`MexUS_targetratediff`: difference between the Mexican central bank's overnight interest rate and the US federal funds rate.

`BdM_any`: a dummy variable equal to one if Mexico's central bank (at the behest of the foreign exchange commission) offered US dollar auctions or dollar futures contracts that day, zero otherwise.

`USpres2016`: a dummy variable equal to one on 9 November 2016, the day after the 8 November 2016 US presidential election, zero otherwise

`NAFTA_roundsandoth`: dichotomous variable equal to one if a NAFTA-related event was being held on that day, zero otherwise.

The authors argue that the Mexican economy is vulnerable to US political events, hence the inclusion of the variables `USpres2016` and `NAFTA_roundsandoth`. Similarly, since the Mexican peso is susceptible to US and Mexican macroeconomic performance and policy shocks, the authors include a number of other macro/financial controls (e.g. `pctchangesp500`, `bondspread10yr_pc...`)

This question is an investigation of GARCH models. The story should make sense and be pretty interesting. Part (a), (b), (d) are standard (ACF/PACF and info criteria, GARCH definition and Q-test on residuals). The rest requires a more careful understanding of these models and pushes students' ability to critically evaluate the results.

- (a) Figure 2 reports the ACF and PACF of the dependent variable and line 29-32 of the do-file report relevant information criteria. Motivating your choice, discuss whether you agree or disagree with the author's choice to model `pctchange_pesot` as an AR(1) process.

Solutions. The choice of an AR(1) model for `pctchange_pesot` is not fully justified by the AC/PACF or information criteria. Students are expected to show that an AR(1) should have a smooth ACF and a PACF with a spike at lag 1 and this is clearly not present in the shown figures. In addition, students should discuss the AIC/BIC tests (using $\ln(\text{SSR})$) and conclude that they also suggest that an AR(1) does not seem appropriate.

- (b) "GARCH models allow us to model both the conditional mean and the conditional error variance as a function of lagged variance, lagged stochastic shocks and exogenous covariates." Define the GARCH(1,1) process estimated by the authors on line 41 of the do-file. In doing so, make sure you explain any assumption needed in the estimation of this model.

Solution. Let \mathbf{x}_t be a vector of the covariates excluding the Tweet dummy above. The model estimated is

$$\text{pctchange_peso}_t = a_0 + a_1 \text{pctchange_peso}_{t-1} + \delta \text{tweetdum}_t + \mathbf{x}'_t \boldsymbol{\gamma} + \epsilon_t \tag{3}$$

$$\epsilon_t = v_t \sqrt{h_t} \tag{4}$$

$$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} \tag{5}$$

We need $\alpha_0 > 0$ and $\alpha_1 + \beta_1 < 1$.

- (c) Consider now the GARCH model estimated on line 53 of the do-file, where the tweet dummy has been included not only in the mean equation for the exchange rate but also in the variance

equation. Compare the unconditional and conditional mean and variance for `pctchange_peso` and ϵ_t under the two different modeling strategies. In other words, what is the role that the additional control variables play in the determination of the statistical properties of the process?

Solution. Note that the text of the question indicates the type of model estimated (even if students do not know the command `het()`). It tells that the tweet dummy is inserted in the mean AND in the conditional variance equation. This is enough to write down the correct model and the difference in the two models is given exactly by the presence of the tweet dummy in the conditional variance equation. The mean equation for `pctchange_pesot` will therefore be the same. This means estimating:

$$\text{pctchange_peso}_t = a_0 + a_1 \text{pctchange_peso}_{t-1} + \delta \text{tweetdum}_t + \mathbf{x}'\boldsymbol{\gamma} + \epsilon_t \quad (6)$$

$$\epsilon_t = v_t \sqrt{h_t} \quad (7)$$

$$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} + \pi \text{tweetdum}_t \quad (8)$$

Students should calculate the conditional and unconditional first two moments of ϵ_t , i.e.:

$$E(\epsilon_t | \epsilon_{t-1}) = E(v_t \sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} + \delta D_t} | \epsilon_{t-1}) = E(v_t) E(\dots) = 0$$

$$E(\epsilon_t) = E(E(\epsilon_t | \epsilon_{t-1})) = E(0) = 0$$

$$E(\epsilon_t^2 | \epsilon_{t-1}) = E_{t-1}(v_t^2 (\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} + \delta D_t)) = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} + \delta D_t$$

$$E(\epsilon_t^2) = E\left(E_{t-1}(\epsilon_t^2)\right) = E(\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} + \delta D_t) = (\alpha_0 + \delta D_t) + \alpha_1 E(\epsilon_{t-1}^2) + \beta_1 E(h_{t-1})$$

Even without full derivation (which we have not produced in class although it is available in the textbook and has been asked in the exam V17, which students had for practice), one can see that the conditional and unconditional variance is affected by the introduction of the dummy just looking at the formulas above. For derivation see the textbook p. 147. Finally, note that an example similar to this is discussed also on p. 155 of the textbook.

- (d) Explain the rationale behind the diagnostics checks on lines 43-49 of the do-file.

Solution. The authors are reporting here standardized errors. From (7) once h_t is estimated correctly we should have that the residuals are white noise as $v_t = \epsilon_t/\sqrt{h_t}$. This is a diagnostic performed using a Q-statistic, at different lags (1, 2, 3). The Q-statistic should be defined (see also exam from H18) and results suggest that the residuals are white noise.

- (e) On line 52 of the do-file I have estimated a model not present in the original file and received an error message, reported on line 40 p. 21 of the log file. Explain: 1) what would be the rationale of including all variables in the variance-equation and 2) which derivatives does Stata refer to.

Solution. A good number of our time series models have been estimated using maximum likelihood. This happens for GARCH models as well. Hence Stata is forming a likelihood function (perfect answers could include its form, see also p. 152) and numerically taking its derivative to find the estimates for the coefficients in the model that maximize this function.

We wrote one in Lecture 5, slide 16. Consider the density for y_t :

$$f(y_t) = \sqrt{\frac{1}{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2\sigma^2}(y_t - \mu)^2 \right\}$$

The likelihood function is the joint density of (y_1, y_2, \dots, y_T) , which, by independence, is the product of the individual density. Stata then maximizes this function, by taking derivatives of it. In a GARCH the variance in this model is not a constant and rather has parameters that needs to be estimated.

The error suggests that it is not possible to find such derivatives. Nonetheless, it would make sense to include all variables as it is not clear a-priori which one should be excluded. In other words, there is no theoretical reason (as far as I know) where we should not expect macro variables to affect not only the average ex rate but also their volatility.

While a perfect answers would include a discussion of the form of the likelihood function as above, a good answer could simply give the intuition as to what Stata is doing.

- (f) Consider the model estimated on line 53 of the do-file (line 41 of the log-file). Did Trumps' tweet affect the volatility of the exchange rate?

Solution. As shown in equation (8) of these solutions, understanding whether Trump's tweet affect volatility requires interpreting the estimates for the parameter π in equation (8). In the conditional variance equation, the tweet dummy is significant at 10% but negative (-1.114756). One could add (unrequested) that in the mean equation the dummy is an economic and statistical zero (0.022). When it comes to the question, at 10% we find an effect, but a negative one!

- (g) Compare and contrast the two models estimated on lines 63-66 of the do-file. Which model would you choose?

Solution. The two models are a GARCH(1,1) and a TGARCH(1,1). The AIC would suggest picking the TGARCH (BIC extremely close), but the negativity constraints are not satisfied in the TGARCH as $\alpha_1 < 0$. It is also a concern in both models that $\alpha_0 < 0$. Overall, if one had to pick, the TGARCH should be disregarded, an EGARCH could be checked but either way these models are not good.

- (h) The authors conclude in their paper that social media usage by the government affect financial markets. Discuss whether the models presented in this question are able to answer this question.

Solution. The model estimated simply includes a dummy for whether something was tweeted, and we find marginal significance of the results – of the sign going in the opposite direction than what expected (reduced volatility). However, here the issue is that some of the results above are deeply worrying (non convergence in the maximum likelihood estimation, issues with signs explained above). Hence, care should be exercised in trusting these conclusions.

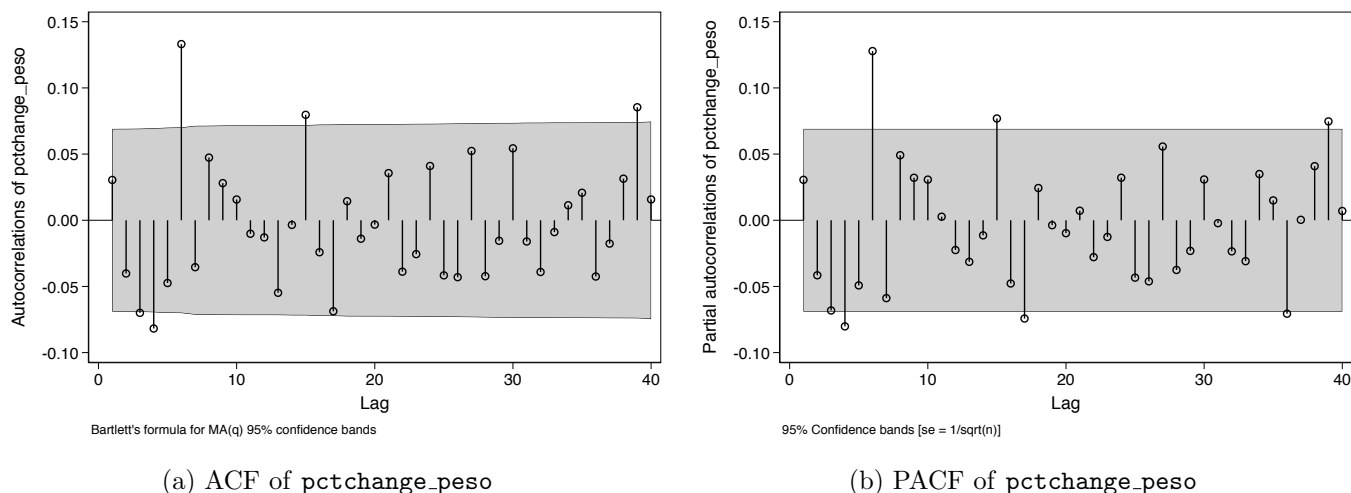


Figure 2: ACF and PACF of pctchange_peso

```

1 *****
2
3 ** This do file performs the analysis discussed
4 ** for QUESTION 1 – Fall 2019. This is based on:
5 ** Aschauer, David Alan. "Is public expenditure productive?."
6 ** Journal of monetary economics 23.2 (1989): 177–200.
7
8 *****
9
10 clear all
11
12 log using "logfile_q1h19.txt", text replace
13
14 use "exam_q1_h19.dta", clear
15
16 tsset year
17
18 *****
19 ** Basic Estimation in Aschauer (1989)
20 *****
21
22 regress yt_kt t nt_kt gt_kt cu
23 regress p t gt_it cu
24
25 *****
26 ** Pre-Testing
27 ** DF critical values:
28 ** 1% = -3.675
29 ** 5% = -2.969
30 ** 10% = -2.617
31 *****
32 //yt
33 reg D.yt_kt L.yt_kt
34 reg D.D.yt_kt D.L.yt_kt
35
36 //nt_kt
37 reg D.nt_kt L.nt_kt
38 reg D.D.nt_kt D.L.nt_kt
39
40 //gt_kt
41 reg D.gt_kt L.gt_kt
42 reg D.D.gt_kt D.L.gt_kt
43
44 //cu
45 reg D.cu L.cu
46 reg D.D.cu D.L.cu
47
48 //p

```

```

49 reg D.p L.p
50 reg D.D.p D.L.p
51
52 //gt_it
53 reg D.gt_it L.gt_it
54 reg D.D.gt_it D.L.gt_it
55
56
57 *****
58 ** VECM
59 *****
60
61 vecrank yt_kt nt_kt gt_kt cu
62
63 vec yt_kt nt_kt gt_kt cu, rank(1)
64
65 veclmar
66
67
68 *****
69 ** Recession
70 *****
71 gen recession=(year == 1974 | year == 1980 | year == 1982 | year ==
1991)
72
73 reg D.nt_kt L.nt_kt recession
74
75
76 log close
77

```

```

-----
name: <unnamed>
log: /Users/co/Documents/Teaching/Courses Taught/Applied Time Series/Exam/logfi
> qlh19.txt
log type: text
opened on: 18 Nov 2019, 17:37:26

```

```

1 .
2 . use "exam_q1_h19.dta", clear

3 .
4 . tsset year
    time variable: year, 1949 to 1985
    delta: 1 unit

```

```

5 .
6 . *****
7 . ** Basic Estimation in Aschauer (1989)
8 . *****
9 .
10 . regress yt_kt t nt_kt gt_kt cu

```

Source	SS	df	MS	Number of obs	=	
-----				F(4, 32)	=	323.33
Model	.093182042	4	.02329551	Prob > F	=	0.0000
Residual	.002305538	32	.000072048	R-squared	=	0.9759
-----				Adj R-squared	=	0.9728
Total	.095487579	36	.002652433	Root MSE	=	.00849

yt_kt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]

t	.0103704	.0018073	5.74	0.000	.0066891 .0140518
nt_kt	.4401808	.0737391	5.97	0.000	.2899791 .5903824
gt_kt	.3787095	.0262646	14.42	0.000	.3252104 .4322087
cu	.4114845	.0365627	11.25	0.000	.3370088 .4859602
_cons	-1.292136	.381811	-3.38	0.002	-2.06986 -.5144127

```

11 . regress p t gt_it cu

```

Source	SS	df	MS	Number of obs	=	
-----				F(3, 33)	=	4358.14
Model	.917211129	3	.305737043	Prob > F	=	0.0000
Residual	.002315051	33	.000070153	R-squared	=	0.9975
-----				Adj R-squared	=	0.9973
Total	.919526179	36	.025542394	Root MSE	=	.00838

p	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
t	.0095906	.0002839	33.78	0.000	.009013	.0101682
gt_it	.3634799	.0161002	22.58	0.000	.3307238	.3962359
cu	.4236126	.0248119	17.07	0.000	.3731323	.4740929
_cons	.774165	.1123446	6.89	0.000	.5455981	1.002732

```

12 .
13 . *****
14 . ** Pre-Testing
15 . ** DF critical values:
16 . ** 1% = -3.675
17 . ** 5% = -2.969
18 . ** 10% = -2.617
19 . *****
20 . //yt
21 . reg D.yt_kt L.yt_kt

```

Source	SS	df	MS	Number of obs	=	36
				F(1, 34)	=	3.73
Model	.004044115	1	.004044115	Prob > F	=	0.0619
Residual	.036884214	34	.00108483	R-squared	=	0.0988
				Adj R-squared	=	0.0723
Total	.040928329	35	.001169381	Root MSE	=	.03294

D.yt_kt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
yt_kt						
L1.	-.2120504	.1098268	-1.93	0.062	-.4352453	.0111444
_cons	.9731052	.504339	1.93	0.062	-.051835	1.998045

```

22 . reg D.D.yt_kt D.L.yt_kt

```

Source	SS	df	MS	Number of obs	=	35
				F(1, 33)	=	30.48
Model	.034839741	1	.034839741	Prob > F	=	0.0000
Residual	.037722758	33	.001143114	R-squared	=	0.4801
				Adj R-squared	=	0.4644
Total	.072562499	34	.002134191	Root MSE	=	.03381

D2.yt_kt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
yt_kt						
LD.	-.9236056	.1672992	-5.52	0.000	-1.263978	-.5832328

_cons	-0.0021094	.0057152	-0.37	0.714	-.0137371	.0095183
-------	------------	----------	-------	-------	-----------	----------

```

23 .
24 . //nt_kt
25 . reg D.nt_kt L.nt_kt

```

Source	SS	df	MS	Number of obs	=	36
				F(1, 34)	=	1.04
Model	.000679914	1	.000679914	Prob > F	=	0.3150
Residual	.022227144	34	.00065374	R-squared	=	0.0297
				Adj R-squared	=	0.0011
Total	.022907058	35	.000654487	Root MSE	=	.02557

D.nt_kt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
nt_kt						
L1.	-.0166279	.0163047	-1.02	0.315	-.049763	.0165072
_cons	.056913	.0794048	0.72	0.478	-.1044569	.2182829

```

26 . reg D.D.nt_kt D.L.nt_kt

```

Source	SS	df	MS	Number of obs	=	35
				F(1, 33)	=	30.21
Model	.020925774	1	.020925774	Prob > F	=	0.0000
Residual	.022855282	33	.000692584	R-squared	=	0.4780
				Adj R-squared	=	0.4621
Total	.043781056	34	.001287678	Root MSE	=	.02632

D2.nt_kt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
nt_kt						
LD.	-.9562642	.1739696	-5.50	0.000	-1.310208	-.6023203
_cons	-.022816	.0061111	-3.73	0.001	-.0352491	-.0103829

```

27 .
28 . //gt_kt
29 . reg D.gt_kt L.gt_kt

```

Source	SS	df	MS	Number of obs	=	36
				F(1, 34)	=	3.58
Model	.000948089	1	.000948089	Prob > F	=	0.0668

Residual		.008992035	34	.000264472	R-squared	=	0.0954
-----					Adj R-squared	=	0.0688
Total		.009940124	35	.000284004	Root MSE	=	.01626

D.gt_kt		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
gt_kt						
l1.		.0679054	.0358649	1.89	0.067	-.0049808 .1407916
_cons		-.3159881	.1646894	-1.92	0.063	-.6506773 .018701

30 . reg D.D.gt_kt D.L.gt_kt

Source		SS	df	MS	Number of obs	=	35
-----					F(1, 33)	=	0.65
Model		.000033271	1	.000033271	Prob > F	=	0.4274
Residual		.001699961	33	.000051514	R-squared	=	0.0192
-----					Adj R-squared	=	-0.0105
Total		.001733231	34	.000050977	Root MSE	=	.00718

D2.gt_kt		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
gt_kt						
LD.		-.0603501	.0750949	-0.80	0.427	-.2131318 .0924317
_cons		-.0006332	.00124	-0.51	0.613	-.0031559 .0018896

31 .
32 . //cu
33 . reg D.cu L.cu

Source		SS	df	MS	Number of obs	=	36
-----					F(1, 34)	=	11.84
Model		.03216652	1	.03216652	Prob > F	=	0.0016
Residual		.09234439	34	.002716011	R-squared	=	0.2583
-----					Adj R-squared	=	0.2365
Total		.124510909	35	.003557455	Root MSE	=	.05212

D.cu		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
cu						
l1.		-.4865097	.1413693	-3.44	0.002	-.7738068 -.1992126
_cons		2.144997	.6229322	3.44	0.002	.8790462 3.410947

Log likelihood = 448.1049 HQIC = -23.64895
 Det(Sigma_ml) = 8.90e-17 SBIC = -22.86329

Equation	Parms	RMSE	R-sq	chi2	P>chi2
D_yt_kt	6	.023793	0.5693	38.33906	0.0000
D_nt_kt	6	.017071	0.8027	117.998	0.0000
D_gt_kt	6	.005167	0.9247	356.1751	0.0000
D_cu	6	.041453	0.5568	36.42666	0.0000

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
<hr/>						
D_yt_kt						
_cel						
L1.	-.3673863	.1596879	-2.30	0.021	-.6803688	-.0544039
yt_kt						
LD.	1.745721	.3793249	4.60	0.000	1.002258	2.489184
nt_kt						
LD.	-1.013841	.3752999	-2.70	0.007	-1.749416	-.2782668
gt_kt						
LD.	.6003356	.4018903	1.49	0.135	-.1873549	1.388026
cu						
LD.	-.5754542	.2292081	-2.51	0.012	-1.024694	-.1262145
_cons	.009935	.0147692	0.67	0.501	-.0190122	.0388822
<hr/>						
D_nt_kt						
_cel						
L1.	-.2945054	.1145764	-2.57	0.010	-.5190711	-.0699397
yt_kt						
LD.	1.697925	.2721666	6.24	0.000	1.164489	2.231362
nt_kt						
LD.	-.4428407	.2692787	-1.64	0.100	-.9706172	.0849358
gt_kt						
LD.	.0195429	.2883573	0.07	0.946	-.545627	.5847129
cu						
LD.	-.691322	.1644574	-4.20	0.000	-1.013653	-.3689914
_cons	-.0063188	.010597	-0.60	0.551	-.0270884	.0144509

D_gt_kt						
_cel						
L1.	.0615324	.0346761	1.77	0.076	-.0064315	.1294963
yt_kt						
LD.	-.1852085	.0823701	-2.25	0.025	-.346651	-.0237661
nt_kt						
LD.	-.1484826	.0814961	-1.82	0.068	-.308212	.0112468
gt_kt						
LD.	.9045266	.0872702	10.36	0.000	.7334802	1.075573
cu						
LD.	.0899933	.0497724	1.81	0.071	-.0075587	.1875454
_cons	-.010049	.0032071	-3.13	0.002	-.0163349	-.0037632

D_cu						
_cel						
L1.	-.3156232	.2782182	-1.13	0.257	-.8609208	.2296744
yt_kt						
LD.	3.681814	.6608835	5.57	0.000	2.386506	4.977122
nt_kt						
LD.	-1.537604	.6538711	-2.35	0.019	-2.819168	-.2560403
gt_kt						
LD.	-.288923	.7001984	-0.41	0.680	-1.661287	1.083441
cu						
LD.	-1.399028	.3993408	-3.50	0.000	-2.181721	-.6163342
_cons	-.0076275	.0257319	-0.30	0.767	-.0580611	.0428061

Cointegrating equations

Equation	Parms	chi2	P>chi2

_cel	3	40.58489	0.0000

Identification: beta is exactly identified

Johansen normalization restriction imposed

beta	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
_ce1						
yt_kt	1
nt_kt	.1200393	.0529118	2.27	0.023	.0163341	.2237445
gt_kt	-.1955167	.1204948	-1.62	0.105	-.4316822	.0406488
cu	-.4634506	.1764092	-2.63	0.009	-.8092062	-.1176951
_cons	-2.145227

44 .
45 . veclmar

Lagrange-multiplier test

lag	chi2	df	Prob > chi2
1	25.6604	16	0.05900
2	17.3955	16	0.36046

H0: no autocorrelation at lag order

46 .
47 .
48 . *****
49 . ** Recession
50 . *****
51 . gen recession=(year == 1974 | year == 1980 | year == 1982 | year == 1991)
52 .
53 . reg D.nt_kt L.nt_kt recession

Source	SS	df	MS	Number of obs	=	36
				F(2, 33)	=	3.95
Model	.004428295	2	.002214148	Prob > F	=	0.0289
Residual	.018478762	33	.000559962	R-squared	=	0.1933
				Adj R-squared	=	0.1444
Total	.022907058	35	.000654487	Root MSE	=	.02366

D.nt_kt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
nt_kt						
L1.	-.028748	.0158004	-1.82	0.078	-.0608942	.0033982
recession	-.0386576	.0149414	-2.59	0.014	-.0690561	-.008259
_cons	.1190749	.077317	1.54	0.133	-.0382277	.2763775

```
54 .
55 .
56 . log close
    name: <unnamed>
    log: /Users/co/Documents/Teaching/Courses Taught/Applied Time Series/Exam/logfi.
> qlh19.txt
    log type: text
    closed on: 18 Nov 2019, 17:37:29
-----
```

```

1 *****
2
3 ** This do file performs the analysis discussed
4 ** for QUESTION 2 – Fall 2019. This is based on:
5 ** Benton, A., and A. Philips. "Does the@realDonaldTrump really
6 ** matter to financial markets.",
7 ** forthcoming American Journal of Political Science (2019).
8 *****
9
10 clear all
11
12 log using "logfile_q2h19.txt", text replace
13
14 global controls "tweetdum l.pctchangesp500 l.bondspread10yr_pc
15 d.lnusdstock d.MexUS_targetratediff BdM_any USpres2016
16 l.USpres2016 NAFTA_roundsandother"
17
18 use "exam_q2_h19.dta", clear
19
20 *****
21
22 ** Figure 2 and Identification mean process
23
24 *****
25
26 ac pctchange_peso
27 pac pctchange_peso
28
29 qui: reg pctchange_peso l.pctchange_peso $controls
30 estat ic
31 qui: reg pctchange_peso $controls
32 estat ic
33
34
35
36 *****
37
38 ** GARCH effects
39
40 *****
41 arch pctchange_peso l.pctchange_peso $controls, arch(1) garch(1)
42 nolag
43
44 predict resids, resid
45 qui predict cond_var, variance

```

```

45 qui gen std_resid = resids/sqrt(cond_var)
46 qui gen std_resid_sq = std_resid^2
47 wntestq std_resid, lags(1)
48 wntestq std_resid, lags(2)
49 wntestq std_resid, lags(3)
50
51
52 arch pctchange_peso l.pctchange_peso $controls, arch(1) garch(1)
het($controls) nolog
53 arch pctchange_peso l.pctchange_peso $controls, arch(1) garch(1)
het(tweetdum) nolog
54
55
56
57 *****
58
59 ** Model comparison
60
61 *****
62
63 arch pctchange_peso l.pctchange_peso $controls, arch(1) garch(1)
het(tweetdum) nolog
64 estat ic
65 arch pctchange_peso l.pctchange_peso $controls, arch(1) garch(1)
het(tweetdum) nolog tarch(1)
66 estat ic
67
68 log close
69

```

.


```

-----
      name: <unnamed>
      log: /Users/co/Documents/Teaching/Courses Taught/Applied Time Series/Exam/logfi
> q2h19.txt
      log type: text
      opened on: 18 Nov 2019, 11:46:00

```

```

1 .
2 . global controls "tweetdum l.pctchangesp500 l.bondspread10yr_pc d.lnUSDstock d.MexUS_t
> etratediff BdM_any USpres2016 l.USpres2016 NAFTA_roundsandother"

```

```

3 .
4 . use "exam_q2_h19.dta", clear

```

```

5 .
6 .
7 .
8 . *****
9 .

```

```

10 . ** Figure 2 and Identification mean process

```

```

11 .
12 . *****

```

```

13 .
14 . ac pctchange_peso

```

```

15 . pac pctchange_peso

```

```

16 .
17 . qui: reg pctchange_peso l.pctchange_peso $controls

```

```

18 . estat ic

```

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	804	-1006.415	-942.7401	11	1907.48	1959.066

Note: N=Obs used in calculating BIC; see [R] BIC note.

```

19 . qui: reg pctchange_peso $controls

```

```

20 . estat ic

```

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
-------	-----	----------	-----------	----	-----	-----

.		804	-1006.415	-943.4223	10	1906.845	1953.741
---	--	-----	-----------	-----------	----	----------	----------

Note: N=Obs used in calculating BIC; see [R] BIC note.

```

21 .
22 .
23 .
24 . *****
25 .
26 . ** GARCH effects
27 .
28 . *****
29 . arch pctchange_peso l.pctchange_peso $controls, arch(1) garch(1) nolog

```

ARCH family regression

```

Sample: 3 - 806                Number of obs   =       804
Distribution: Gaussian          Wald chi2(10)  =       15.56
Log likelihood = -929.726      Prob > chi2    =       0.1130

```

pctchange_peso	Coef.	OPG		z	P> z	[95% Conf. Interval]	
		Std. Err.					
pctchange_peso							
pctchange_peso							
L1.	.024351	.0340445		0.72	0.474	-.042375	.091077
tweetdum	.0132815	.0580325		0.23	0.819	-.10046	.1270231
pctchangesp500							
L1.	.0559377	.0395168		1.42	0.157	-.0215138	.1333893
bondspread10yr_pc							
L1.	.0350685	.0248522		1.41	0.158	-.0136409	.0837778
lnusdstock							
D1.	10.61515	16.10443		0.66	0.510	-20.94894	42.17924
MexUS_targetratediff							
D1.	-1.272865	.520536		-2.45	0.014	-2.293097	-.2526332
BdM_any	.1193477	.0598129		2.00	0.046	.0021166	.2365788
USpres2016							
--.	-1.512456	84818.18		-0.00	1.000	-166242.1	166239.1
L1.	8.52483	7816.193		0.00	0.999	-15310.93	15327.98

NAFTA_roundsandoth		.0932701	.1397222	0.67	0.504	-.1805804	.3671206
_cons		-.0186346	.0358608	-0.52	0.603	-.0889205	.0516513

ARCH							
arch							
L1.		.0769638	.0270614	2.84	0.004	.0239245	.1300031
garch							
L1.		.8450199	.0656279	12.88	0.000	.7163916	.9736482
_cons		.0484788	.0290683	1.67	0.095	-.008494	.1054516

```

30 .
31 . predict resid, resid
    (2 missing values generated)

32 . qui predict cond_var, variance

33 . qui gen std_resid = resid/sqrt(cond_var)

34 . qui gen std_resid_sq = std_resid^2

35 . wntestq std_resid, lags(1)

Portmanteau test for white noise
-----
Portmanteau (Q) statistic =    0.4525
Prob > chi2(1)           =    0.5012

36 . wntestq std_resid, lags(2)

Portmanteau test for white noise
-----
Portmanteau (Q) statistic =    0.4544
Prob > chi2(2)           =    0.7968

37 . wntestq std_resid, lags(3)

Portmanteau test for white noise
-----
Portmanteau (Q) statistic =    1.7508
Prob > chi2(3)           =    0.6257

38 .
39 .
40 . arch pctchange_peso l.pctchange_peso $controls, arch(1) garch(1) het($controls) nolog
    numerical derivatives are approximate
    flat or discontinuous region encountered

```

```
41 . arch pctchange_peso l.pctchange_peso $controls, arch(1) garch(1) het(tweetdum) nolog
```

```
ARCH family regression -- multiplicative heteroskedasticity
```

```
Sample: 3 - 806                    Number of obs   =          804
Distribution: Gaussian               Wald chi2(10)   =          14.36
Log likelihood = -928.0875          Prob > chi2     =          0.1573
```

		OPG			[95% Conf. Interval]	
pctchange_peso	Coef.	Std. Err.	z	P> z		

pctchange_peso						
pctchange_peso						
L1.	.0229936	.0331503	0.69	0.488	-.0419799	.0879671
tweetdum	.0224462	.0565912	0.40	0.692	-.0884704	.1333628
pctchangesp500						
L1.	.0572399	.0394856	1.45	0.147	-.0201504	.1346302
bondspread10yr_pc						
L1.	.0367846	.0246838	1.49	0.136	-.0115949	.085164
lnusdstock						
D1.	10.01577	16.26289	0.62	0.538	-21.85891	41.89045
MexUS_targetratediff						
D1.	-1.247495	.5378157	-2.32	0.020	-2.301594	-.1933957
BdM_any	.1141022	.0595931	1.91	0.056	-.0026982	.2309026
USpres2016						
--.	-1.513976	2604.896	-0.00	1.000	-5107.016	5103.988
L1.	8.527756	1134.162	0.01	0.994	-2214.39	2231.445
NAFTA_roundsandother						
_cons	.0798833	.13139	0.61	0.543	-.1776363	.337403
	-.0196001	.0373185	-0.53	0.599	-.092743	.0535428

HET						
tweetdum	-1.114756	.6579774	-1.69	0.090	-2.404368	.1748556
_cons	-2.814345	.4990637	-5.64	0.000	-3.792492	-1.836198

ARCH						
arch						
L1.	.0688467	.0244841	2.81	0.005	.0208587	.1168346
garch						

L1. | .8544119 .0550155 15.53 0.000 .7465835 .9622404

```

42 .
43 .
44 .
45 . *****
46 .
47 . ** Model comparison
48 .
49 . *****
50 .
51 . arch pctchange_peso l.pctchange_peso $controls, arch(1) garch(1) het(tweetdum) nolog

```

ARCH family regression -- multiplicative heteroskedasticity

```

Sample: 3 - 806                Number of obs   =      804
Distribution: Gaussian          Wald chi2(10)   =     14.36
Log likelihood = -928.0875      Prob > chi2     =     0.1573

```

		OPG				[95% Conf. Interval]	
pctchange_peso		Coef.	Std. Err.	z	P> z		
pctchange_peso							
	pctchange_peso						
	L1.	.0229936	.0331503	0.69	0.488	-.0419799	.0879671
	tweetdum	.0224462	.0565912	0.40	0.692	-.0884704	.1333628
pctchangesp500							
	L1.	.0572399	.0394856	1.45	0.147	-.0201504	.1346302
bondspread10yr_pc							
	L1.	.0367846	.0246838	1.49	0.136	-.0115949	.085164
lnusdstock							
	D1.	10.01577	16.26289	0.62	0.538	-21.85891	41.89045
MexUS_targetratediff							
	D1.	-1.247495	.5378157	-2.32	0.020	-2.301594	-.1933957
BdM_any							
	L1.	.1141022	.0595931	1.91	0.056	-.0026982	.2309026
USpres2016							
	--.	-1.513976	2604.896	-0.00	1.000	-5107.016	5103.988
	L1.	8.527756	1134.162	0.01	0.994	-2214.39	2231.445
NAFTA_roundsandother							
	L1.	.0798833	.13139	0.61	0.543	-.1776363	.337403

	_cons	-.0196001	.0373185	-0.53	0.599	-.092743	.0535428

HET	tweetdum	-1.114756	.6579774	-1.69	0.090	-2.404368	.1748556
	_cons	-2.814345	.4990637	-5.64	0.000	-3.792492	-1.836198

ARCH	arch						
	L1.	.0688467	.0244841	2.81	0.005	.0208587	.1168346
	garch						
	L1.	.8544119	.0550155	15.53	0.000	.7465835	.9622404

52 . estat ic

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	804	.	-928.0875	15	1886.175	1956.519

Note: N=Obs used in calculating BIC; see [R] BIC note.

53 . arch pctchange_peso l.pctchange_peso \$controls, arch(1) garch(1) het(tweetdum) nolog
> ch(1)

ARCH family regression -- multiplicative heteroskedasticity

Sample: 3 - 806 Number of obs = 804
Distribution: Gaussian Wald chi2(10) = 21.80
Log likelihood = -924.8017 Prob > chi2 = 0.0162

	Coef.	OPG Std. Err.	z	P> z	[95% Conf. Interval]
pctchange_peso					
pctchange_peso					
L1.	.0343943	.0317941	1.08	0.279	-.027921 .0967096
tweetdum	-.006067	.0599344	-0.10	0.919	-.1235363 .1114023
pctchangesp500					
L1.	.0681428	.0358083	1.90	0.057	-.0020403 .1383259
bondspread10yr_pc					
L1.	.0276071	.025643	1.08	0.282	-.0226523 .0778665

	lnusdstock						
	D1.	10.34933	15.58921	0.66	0.507	-20.20495	40.90361
MexUS_targetratediff							
	D1.	-1.214277	.5098728	-2.38	0.017	-2.213609	-.2149443
	BdM_any	.1366726	.0515383	2.65	0.008	.0356593	.2376858
	USpres2016						
	--.	-1.535542	1434.567	-0.00	0.999	-2813.236	2810.165
	L1.	8.516001	740.3987	0.01	0.991	-1442.639	1459.671
NAFTA_roundsandother		.2529498	.1322614	1.91	0.056	-.0062778	.5121774
	_cons	-.0169625	.0392824	-0.43	0.666	-.0939545	.0600295

HET							
	tweetdum	-1.557858	3.984024	-0.39	0.696	-9.366401	6.250686
	_cons	-5.00012	.6372736	-7.85	0.000	-6.249153	-3.751087

ARCH							
	arch						
	L1.	-.0259172	.0093108	-2.78	0.005	-.0441662	-.0076683
	tarch						
	L1.	.05606	.0108793	5.15	0.000	.0347369	.077383
	garch						
	L1.	.9889193	.0081029	122.05	0.000	.9730379	1.004801

54 . estat ic

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	804	.	-924.8017	16	1881.603	1956.637

Note: N=Obs used in calculating BIC; see [R] BIC note.

55 .

56 . log close

```

name: <unnamed>
log: /Users/co/Documents/Teaching/Courses Taught/Applied Time Series/Exam/logfi.
> q2h19.txt
log type: text
closed on: 18 Nov 2019, 11:48:01

```