

AQM Assignment One

Ada Žiemyte, Shalva Kilasonia, Anders Riis

March 2013, College of Europe

Contents

a. ARIMA Models	1
a.i. ARIMA model for ln_inv	1
a.ii. ARIMA model for ln_inc	6
a.iii. ARIMA model for $ln_consump$	12
b. Dynamic Linear Model	14
c. Error Correction Model	16
d. Model Specification	18
e. Comparison of Models	20
f. VAR(4) Model	20
g. Final Model	27

a. ARIMA Models

Identify, estimate and diagnostic check ARIMA models for each of the following variables: ln_inv , ln_inc , and $ln_consump$.

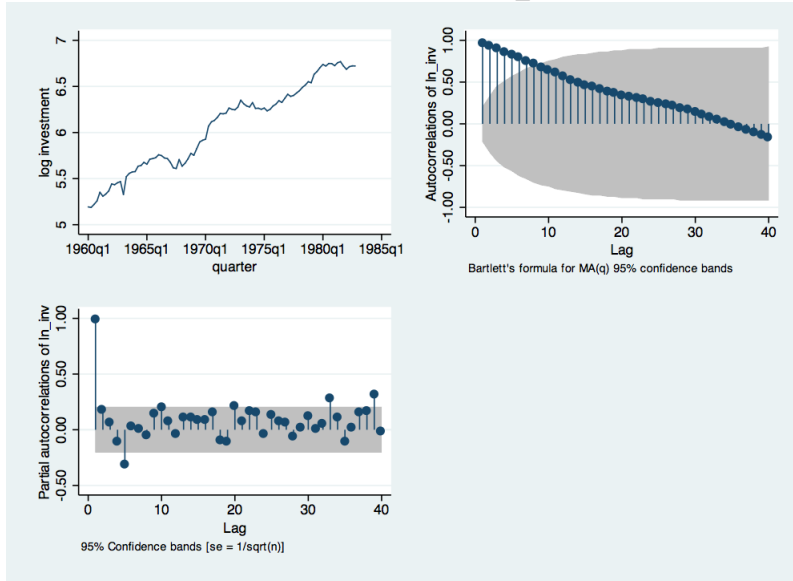
To identify an ARIMA model we must check for stationarity, correct for stationarity and examine the graphs for the autocorrelations and partial autocorrelations.¹

a.i. ARIMA model for ln_inv

As can be seen in figure 1 on the following page, the ACF decays monotonically. That the ACF decays slowly indicates that the data is non-stationary, i.e. that

¹<http://www.polsci.wvu.edu/duval/ps791c/Notes/Stata/arimaident.html>

Figure 1: Graphs for \ln_inv



we have a non-stationary mean.² Stationarity is when we have a flat looking series without trend.³ The scatterplot in figure 1 also indicates that the series is non-stationary as we can see a clear trend. The PACF cuts off after one lag (see figure 1), so we probably have an AR(1) process. However, because of non-stationarity we first attempt to correct for this by taking the first difference.

Now in figure 2 on the next page we see that the scatterplot is flatter, portraying stationarity. The ACF decays in an alternating or oscillating fashion, probably portraying some sort of MA as it seems to cut off (becomes zero) after four values or after zero values ($q = 0$). The PACF never seems to cut off ($p = 0$), although it can be seen as decaying. Thus, we have evidence of an ARIMA(0,1,0) model:

```
. arima ln_inv, arima(0,1,0)

(setting optimization to BHHH)
Iteration 0:  log likelihood = 154.00497
Iteration 1:  log likelihood = 154.00497

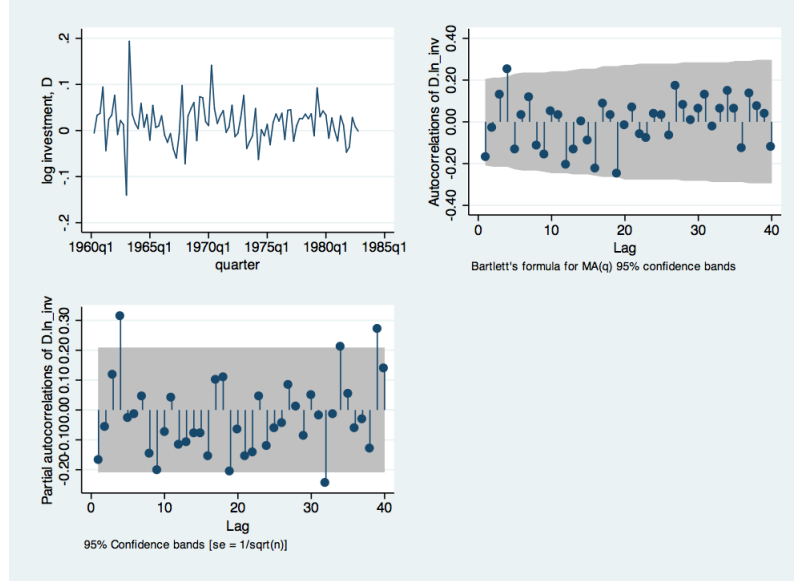
ARIMA regression

Sample: 1960q2 - 1982q4                Number of obs   =      91
                                         Wald chi2(.)    =      .
Log likelihood = 154.005                 Prob > chi2     =      .
```

²<http://www.polsci.wvu.edu/duval/ps791c/Notes/Stata/arimaident.html>

³<http://www.itl.nist.gov/div898/handbook/pmc/section4/pmc442.htm>

Figure 2: Graphs for \ln_inv after correcting for stationarity



		OPG				
D.ln_inv		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
ln_inv						
	_cons	.0167964	.0047276	3.55	0.000	.0075304 .0260623
	/sigma	.044543	.0020718	21.50	0.000	.0404823 .0486037

Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at z

It does not make sense to do an `armaroot` test as we have no AR or MA parameters in the model.

Another way of interpreting the graphs in figure 2 is that we have an ARIMA(4,1,4) model, the results of which are:

```
. arima ln_inv, arima(4,1,4)

(setting optimization to BHHH)
Iteration 0:  log likelihood = 151.19416
Iteration 1:  log likelihood = 153.8349
Iteration 2:  log likelihood = 155.72441
Iteration 3:  log likelihood = 158.06115
Iteration 4:  log likelihood = 158.78522
(switiching optimization to BFGS)
Iteration 5:  log likelihood = 160.42741
Iteration 6:  log likelihood = 161.31532
Iteration 7:  log likelihood = 161.57824
Iteration 8:  log likelihood = 162.63005
Iteration 9:  log likelihood = 162.99665
```

```

Iteration 10: log likelihood = 163.47528
Iteration 11: log likelihood = 163.81731
Iteration 12: log likelihood = 164.09801
Iteration 13: log likelihood = 164.17117
Iteration 14: log likelihood = 164.23693
(switching optimization to BHHH)
Iteration 15: log likelihood = 164.28633
Iteration 16: log likelihood = 164.28633 (backed up)
Iteration 17: log likelihood = 164.28633 (not concave)
Iteration 18: log likelihood = 164.29517
Iteration 19: log likelihood = 164.29517 (backed up)
(switching optimization to BFGS)
Iteration 20: log likelihood = 164.29519 (backed up)
Iteration 21: log likelihood = 164.29682
Iteration 22: log likelihood = 164.30153
Iteration 23: log likelihood = 164.30185
Iteration 24: log likelihood = 164.3047
Iteration 25: log likelihood = 164.30508
Iteration 26: log likelihood = 164.30518
Iteration 27: log likelihood = 164.30518

```

ARIMA regression

```

Sample: 1960q2 - 1982q4      Number of obs      =      91
                             Wald chi2(7)           =     361.32
Log likelihood = 164.3052    Prob > chi2        =     0.0000

```

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

ln_inv							
_cons		.0170483	.0066675	2.56	0.011	.0039803	.0301164

ARMA							
ar							
L1.		.250352	.3175816	0.79	0.431	-.3720964	.8728004
L2.		-.6410894	.3055835	-2.10	0.036	-1.240022	-.0421568
L3.		.6959524	.2326489	2.99	0.003	.2399689	1.151936
L4.		-.1234277	.371418	-0.33	0.740	-.8513935	.6045381
ma							
L1.		-.4332859	.3581174	-1.21	0.226	-1.135183	.2686114
L2.		.7828925	.3674864	2.13	0.033	.0626323	1.503153
L3.		-.7487543
L4.		.5336774	.347815	1.53	0.125	-.1480275	1.215382

/sigma		.0387658	.0050688	7.65	0.000	.0288311	.0487004

Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at z

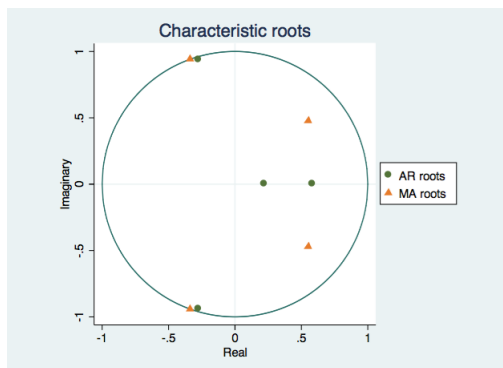
To test for stationarity we use the -armaroots- command to know whether the absolute value of the roots of the characteristic equation are smaller than one, which is required for stationarity. As we can see in the following table all the roots are smaller than one in absolute value:

```
. armaroots
```

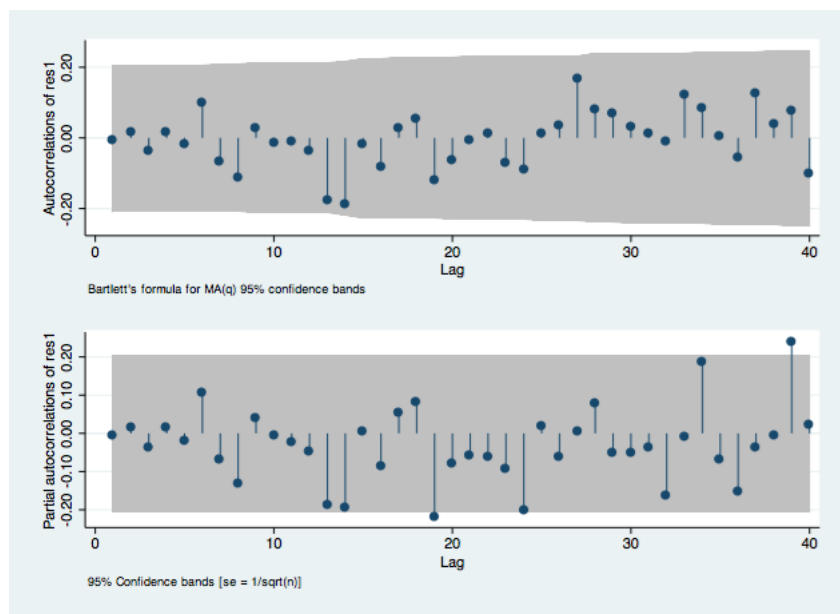
Characteristic roots of AR-polynomial			
Characteristic roots	Modulus	Period	
-.2762973 + .9377075i	.977566	3.38	
-.2762973 - .9377075i	.977566	3.38	
.5804232	.580423		
.2225233	.222523		

Characteristic roots of MA-polynomial			
Characteristic roots	Modulus	Period	
-.3382539 + .941054i	.999999	3.28	
-.3382539 - .941054i	.999999	3.28	
.5548968 + .4751503i	.730533	8.87	
.5548968 - .4751503i	.730533	8.87	

Graphically we can also confirm that none of the roots are larger than one in absolute value:



We now turn to the ACF and the PACF for the residuals to confirm that these have no distinct pattern:



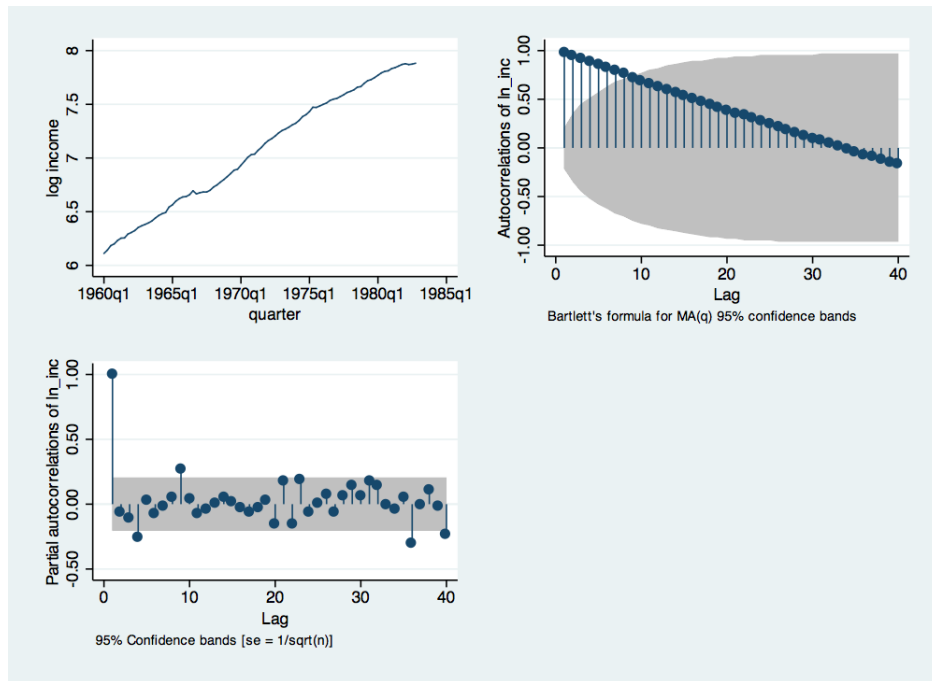
The graph above indeed seems to confirm that the residuals are white noise.
The Portmanteau statistics confirms this by yielding:

```
. wntestq res1

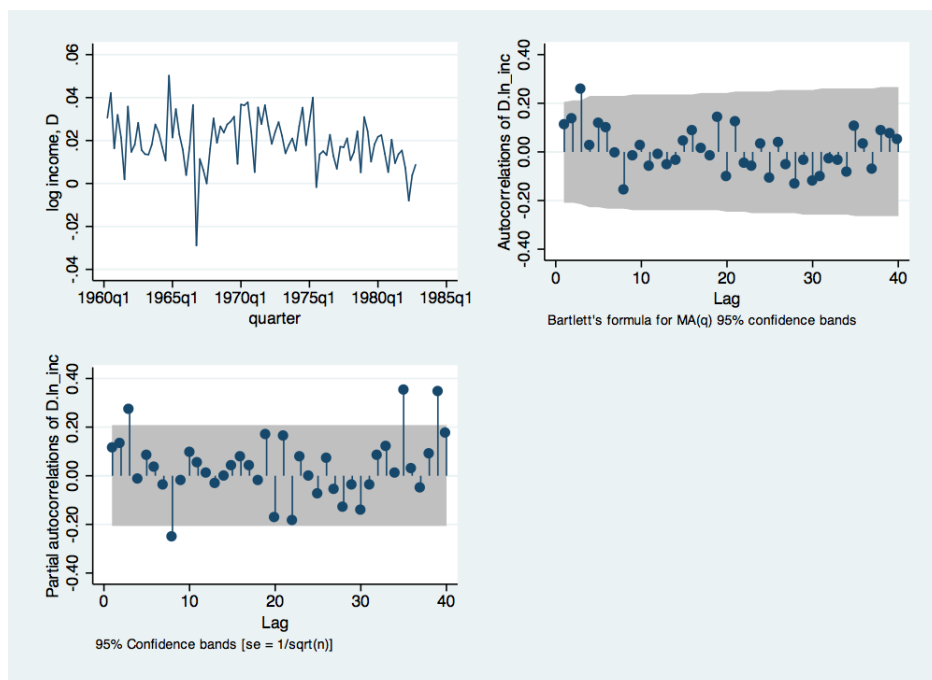
Portmanteau test for white noise
-----
Portmanteau (Q) statistic =    30.0852
Prob > chi2(40)           =    0.8728
```

a.ii. ARIMA model for \ln_inc

We now turn to the variable \ln_inc .



The graphs shown above again demonstrate that we probably have non-stationary data, so we do a first difference transformation again.



The results of this first difference transformation can be seen in the graphs above. Now the data does not portray a non-stationary mean, so we decide that $d = 1$. Again the case is more ambiguous for the ACF and PACF. We could again conclude that the ACF never cuts off, so we have a $MA(0)$, where $q = 0$, and the PACF never cuts off (becomes zero), so $p = 0$. Estimating an $ARIMA(0,1,0)$ model gives the results:

```
. arima ln_inc, arima(0,1,0)

(setting optimization to BHHH)
Iteration 0:  log likelihood = 274.34425
Iteration 1:  log likelihood = 274.34425

ARIMA regression

Sample: 1960q2 - 1982q4                Number of obs   =      91
                                         Wald chi2(.)     =      .
Log likelihood = 274.3443                Prob > chi2      =      .

-----+-----
      |               OPG
D.ln_inc |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
ln_inc  |
   _cons |      .019464   .0012907    15.08   0.000     .0169343     .0219938
-----+-----
      /sigma |      .0118704   .0006448    18.41   0.000     .0106066     .0131342
```



```

-----
Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at z
(1 missing value generated)
. armaroots
No AR or MA parameters in model

```

Another possibility is however to interpret the ACF and the PACF cuts off after three after significant values so we have an ARIMA(3,1,3) model:

```

. arima ln_inc, arima(3,1,3)

(setting optimization to BHHH)
Iteration 0: log likelihood = 270.31907
Iteration 1: log likelihood = 274.68144
Iteration 2: log likelihood = 276.35436
Iteration 3: log likelihood = 277.54725
Iteration 4: log likelihood = 278.059
(swimming optimization to BFGS)
Iteration 5: log likelihood = 278.30896
Iteration 6: log likelihood = 278.77259
Iteration 7: log likelihood = 278.78656
Iteration 8: log likelihood = 279.11859
Iteration 9: log likelihood = 279.5552
Iteration 10: log likelihood = 279.5552 (backed up)
Iteration 11: log likelihood = 279.70686 (backed up)
Iteration 12: log likelihood = 280.14331
Iteration 13: log likelihood = 280.39292
Iteration 14: log likelihood = 280.4006
(swimming optimization to BHHH)
Iteration 15: log likelihood = 280.40117
Iteration 16: log likelihood = 280.40177 (backed up)
Iteration 17: log likelihood = 280.4201
Iteration 18: log likelihood = 280.42064
Iteration 19: log likelihood = 280.43726
(swimming optimization to BFGS)
Iteration 20: log likelihood = 280.49572
Iteration 21: log likelihood = 280.67651
Iteration 22: log likelihood = 280.74026
Iteration 23: log likelihood = 280.7466
Iteration 24: log likelihood = 281.03547
Iteration 25: log likelihood = 281.05958
Iteration 26: log likelihood = 281.06647
Iteration 27: log likelihood = 281.06818
Iteration 28: log likelihood = 281.06845
Iteration 29: log likelihood = 281.06846

```

ARIMA regression

```

Sample: 1960q2 - 1982q4      Number of obs      =      91
                             Wald chi2(6)           =      47.81
Log likelihood = 281.0685    Prob > chi2       =      0.0000

```

```

-----
              |               OPG
          D.ln_inc |      Coef.   Std. Err.   z   P>|z|   [95% Conf. Interval]
-----+-----

```

```

ln_inc |
      _cons |      .0193814      .0022928      8.45      0.000      .0148877      .0238752
-----+-----
ARMA |
      ar |
      L1. |      .3524409      .3992522      0.88      0.377      -.4300791      1.134961
      L2. |      .5973378      .1728554      3.46      0.001      .2585474      .9361281
      L3. |      -.3721901      .3633335      -1.02      0.306      -1.084311      .3399305
      ma |
      L1. |      -.2371601      480.2177      -0.00      1.000      -941.4465      940.9722
      L2. |      -.6004014      593.9994      -0.00      0.999      -1164.818      1163.617
      L3. |      .6367674      305.7404      0.00      0.998      -598.6033      599.8769
-----+-----
      /sigma |      .0108757      2.61138      0.00      0.498      0      5.129086
-----+-----

```

Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at z (1 missing value generated)

```
. armaroots
```

```

      Characteristic roots of AR-polynomial
+-----+-----+
| Characteristic roots | Modulus | Period |
+-----+-----+
| -.8551112            | .855111 |        |
| .603776 + .2659097i | .659737 | 15.1   |
| .603776 - .2659097i | .659737 | 15.1   |
+-----+-----+

```

```

      Characteristic roots of MA-polynomial
+-----+-----+
| Characteristic roots | Modulus | Period |
+-----+-----+
| -1.000003           | 1        |        |
| .6185815 + .5041055i | .797976 | 9.19   |
| .6185815 - .5041055i | .797976 | 9.19   |
+-----+-----+

```

All the significance values for moving average is however insignificant and the characteristic roots of the MA-polynomial is not smaller than one in absolute value, so we postulate an ARIMA(3,1,0):

```

. arima ln_inc, arima(3,1,0)

(setting optimization to BHHH)
Iteration 0:  log likelihood = 278.89994
Iteration 1:  log likelihood = 278.90832
Iteration 2:  log likelihood = 278.91024
Iteration 3:  log likelihood = 278.91141
Iteration 4:  log likelihood = 278.91196
(switiching optimization to BFGS)
Iteration 5:  log likelihood = 278.91231
Iteration 6:  log likelihood = 278.91275
Iteration 7:  log likelihood = 278.91278
Iteration 8:  log likelihood = 278.91278

```

ARIMA regression

```

Sample: 1960q2 - 1982q4      Number of obs   =      91
                             Wald chi2(3)      =       7.42
Log likelihood = 278.9128     Prob > chi2      =     0.0597

```

		OPG		z	P> z	[95% Conf. Interval]	
D.ln_inc	Coef.	Std. Err.					
ln_inc							
_cons	.0193577	.0022504	8.60	0.000	.014947	.0237684	
ARMA							
ar							
L1.	.0680786	.1137906	0.60	0.550	-.1549469	.291104	
L2.	.1220981	.1069794	1.14	0.254	-.0875776	.3317738	
L3.	.2737841	.1216749	2.25	0.024	.0353056	.5122626	
/sigma	.0112707	.0007352	15.33	0.000	.0098298	.0127116	

Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at z (1 missing value generated)

```
. armaroots
```

Characteristic roots of AR-polynomial

Characteristic roots	Modulus	Period
.7373076	.737308	
-.3346145 + .5092766i	.609368	2.92
-.3346145 - .5092766i	.609368	2.92

```
. wntestq res1
```

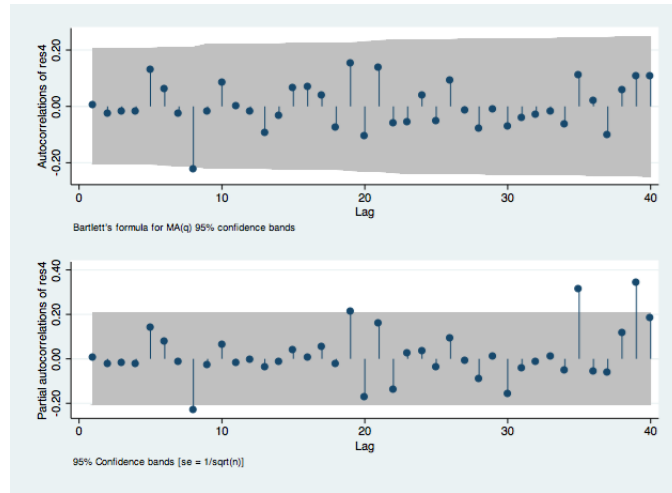
Portmanteau test for white noise

```

Portmanteau (Q) statistic = 30.2217
Prob > chi2(40)           = 0.8689

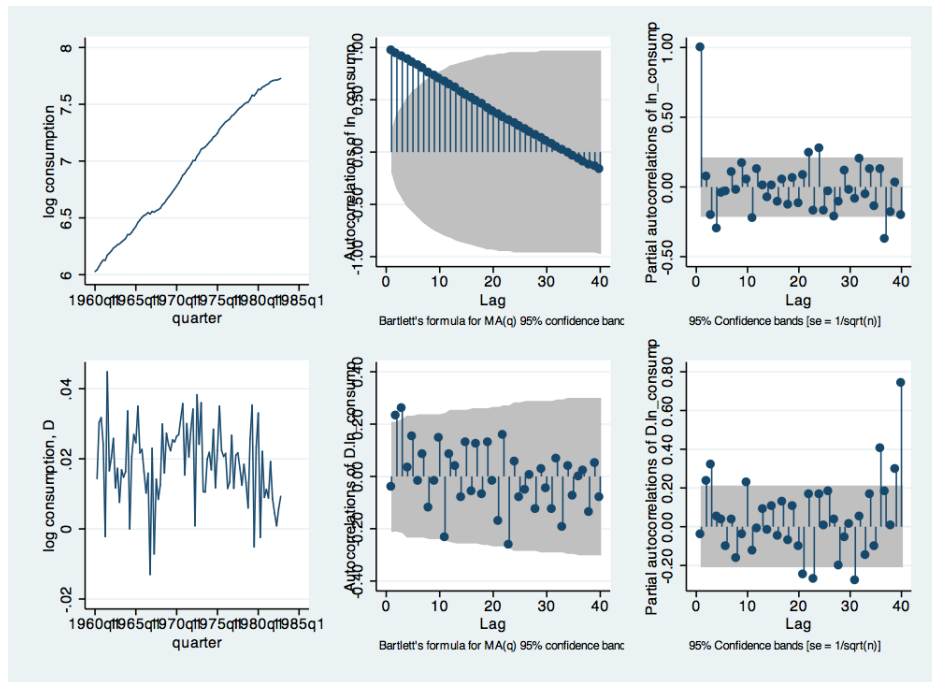
```

The Portmanteau test above indicates that the residuals are white noise. To verify this, we also produce the residuals of the ARIMA(3,1,0) model:



a.iii. ARIMA model for $\ln_consump$

Identifying an ARIMA model for the variable $\ln_consump$ we again first look at some graphs. Again we have to take the first difference to correct for stationarity. After this is done we see in the lower pane that we have a stationary mean:



We postulate an ARIMA(2,1,2):

```
. arima ln_inc, arima(2,1,2)

(setting optimization to BHHH)
Iteration 0: log likelihood = 270.80789
Iteration 1: log likelihood = 273.46539
Iteration 2: log likelihood = 274.91076
Iteration 3: log likelihood = 276.40311
Iteration 4: log likelihood = 276.40314 (backed up)
(switching optimization to BFGS)
Iteration 5: log likelihood = 276.40314 (backed up)
Iteration 6: log likelihood = 276.88434
Iteration 7: log likelihood = 277.73776
Iteration 8: log likelihood = 278.36479
Iteration 9: log likelihood = 278.46702
Iteration 10: log likelihood = 278.47605
Iteration 11: log likelihood = 278.47636
Iteration 12: log likelihood = 278.47641
Iteration 13: log likelihood = 278.47642

ARIMA regression

Sample: 1960q2 - 1982q4                Number of obs   =          91
                                         Wald chi2(4)    =          81.47
Log likelihood = 278.4764                Prob > chi2     =          0.0000
```

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

ln_inc							
	_cons	.0192823	.0026045	7.40	0.000	.0141776	.0243869

ARMA							
	ar						
	L1.	-.0365622	.2870042	-0.13	0.899	-.59908	.5259557
	L2.	.7588994	.2752277	2.76	0.006	.2194631	1.298336
	ma						
	L1.	.2707034	207.4331	0.00	0.999	-406.2907	406.8321
	L2.	-.7292833	151.3058	-0.00	0.996	-297.2832	295.8246

	/sigma	.0112218	1.163677	0.01	0.496	0	2.291987

Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at z (1 missing value generated)

```
. armaroots
```

Characteristic roots of AR-polynomial

+-----+			
Characteristic roots	Modulus	Period	
+-----+			
-.8896212	.889621		
.853059	.853059		
+-----+			

Characteristic roots of MA-polynomial

```

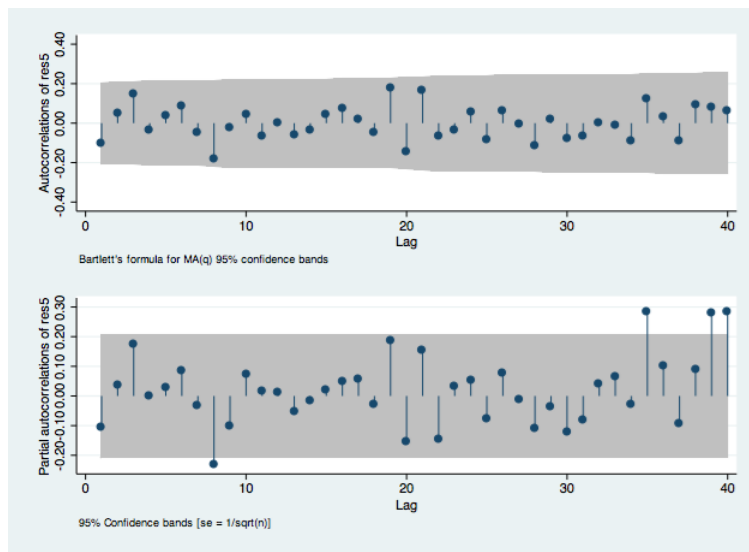
+-----+
| Characteristic roots | Modulus | Period |
+-----+-----+
| -.9999923          | .999992 |        |
| .7292889           | .729289 |        |
+-----+-----+

. wntestq res1

Portmanteau test for white noise
-----
Portmanteau (Q) statistic =    33.7915
Prob > chi2(40)          =    0.7447

```

The ACF and the PACF of the residuals is given by:



b. Dynamic Linear Model

Formulate and estimate a dynamic linear model with four lags for $\ln_consump$ with \ln_inc and \ln_inv as explanatory variables.

The model can be written as

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + a_4 y_{t-4} + b_0 \ln_inc_t + c_0 \ln_inv_t + u_t, \quad (1)$$

where y is $\ln_consump$.⁴

The results of a dynamic linear model with four lags for $\ln_consump$ are:

⁴See AQM lecture notes chapter 3, p. 18.

```
. reg ln_consump L(1/4).ln_consump ln_inc ln_inv
```

Source	SS	df	MS	Number of obs = 88		
Model	22.3174282	6	3.71957137	F(6, 81) =55970.25		
Residual	.005382954	81	.000066456	Prob > F = 0.0000		
Total	22.3228112	87	.256584037	R-squared = 0.9998		
				Adj R-squared = 0.9997		
				Root MSE = .00815		

ln_consump	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ln_consump						
L1.	.4046597	.1056179	3.83	0.000	.1945131	.6148062
L2.	.2783657	.1128595	2.47	0.016	.0538107	.5029207
L3.	.1180614	.111389	1.06	0.292	-.1035678	.3396905
L4.	-.1543535	.0870547	-1.77	0.080	-.327565	.0188579
ln_inc	.3326671	.0489638	6.79	0.000	.2352445	.4300897
ln_inv	.009699	.0120582	0.80	0.424	-.0142931	.0336911
_cons	.0492794	.0134815	3.66	0.000	.0224554	.0761033

Mis-specification tests are:

Ramsey RESET test using powers of the fitted values of ln_consump

Ho: model has no omitted variables

F(3, 78) = 6.02

Prob > F = 0.0010

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance

Variables: fitted values of ln_consump

chi2(1) = 1.10

Prob > chi2 = 0.2944

White's test for Ho: homoskedasticity

against Ha: unrestricted heteroskedasticity

chi2(27) = 30.03

Prob > chi2 = 0.3130

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	30.03	27	0.3130
Skewness	6.37	6	0.3831
Kurtosis	4.75	1	0.0294
Total	41.14	34	0.1863

(n = 88)	D-H	P-value	asy.	P-value
Residuals	6.1568	0.0460	3.6067	0.1647

Breusch-Godfrey LM test for autocorrelation

lags(p)	chi2	df	Prob > chi2
1	2.096	1	0.1476

H0: no serial correlation

LM test for autoregressive conditional heteroskedasticity (ARCH)

lags(p)	chi2	df	Prob > chi2
1	1.558	1	0.2120

H0: no ARCH effects vs. H1: ARCH(p) disturbance

c. Error Correction Model

Reformulate the model in b. as an error correction model; store the log-likelihood value using the command `-estimates store llu-`; and subject the model to mis-specification testing.

Writing equation 1 on page 14 as an error correction model yields:

$$\Delta y_t = a_0 + \alpha_1 \Delta y_{t-1} + \alpha_2 \Delta y_{t-2} + \alpha_3 y_{t-3} + \beta_0 \Delta \ln_inc_t + \gamma_0 \Delta \ln_inv_t + a y_{t-1} + u_t,$$

again where y is $\ln_consump$.⁵ Running the regression gives:

Source	SS	df	MS	Number of obs =	88
Model	.004719858	6	.000786643	F(6, 81) =	11.06
Residual	.005759951	81	.000071111	Prob > F =	0.0000
Total	.010479809	87	.000120458	R-squared =	0.4504
				Adj R-squared =	0.4097
				Root MSE =	.00843

dln_consump	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
L1.	-.2673194	.087738	-3.05	0.003	-.4418905 -.0927483
L2.	.1660194	.0843342	1.97	0.052	-.0017792 .333818
L3.	.2591583	.0865452	2.99	0.004	.0869605 .4313562
dln_inc	.4554999	.0815985	5.58	0.000	.2931446 .6178553
dln_inv	.062451	.0202667	3.08	0.003	.0221265 .1027754

⁵See answer to assignment 4, IQM, p. 13.

ln_consump							
L1.		-.0004039	.0018446	-0.22	0.827	-.0040741	.0032664
_cons		.008395	.0138116	0.61	0.545	-.0190858	.0358757

The results of mis-specification tests are:

Ramsey RESET test using powers of the fitted values of dln_consump

Ho: model has no omitted variables

F(3, 78) = 1.19

Prob > F = 0.3193

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance

Variables: fitted values of dln_consump

chi2(1) = 0.47

Prob > chi2 = 0.4911

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance

Variables: L.dln_consump L2.dln_consump L3.dln_consump dln_inc dln_inv L.ln_consump

chi2(6) = 7.60

Prob > chi2 = 0.2686

White's test for Ho: homoskedasticity

against Ha: unrestricted heteroskedasticity

chi2(27) = 38.39

Prob > chi2 = 0.0719

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	38.39	27	0.0719
Skewness	9.38	6	0.1532
Kurtosis	0.65	1	0.4218
Total	48.42	34	0.0519

(n = 88)	D-H	P-value	asy.	P-value
Residuals	2.5948	0.2732	2.4639	0.2917

Breusch-Godfrey LM test for autocorrelation

lags(p)	chi2	df	Prob > chi2
---------	------	----	-------------

1		4.266	1	0.0389
2		8.164	2	0.0169
3		9.789	3	0.0205
4		10.731	4	0.0298

H0: no serial correlation
LM test for autoregressive conditional heteroskedasticity (ARCH)

lags(p)		chi2	df	Prob > chi2
1		3.184	1	0.0743
2		8.175	2	0.0168
3		1.331	3	0.7217
4		0.949	4	0.9175

H0: no ARCH effects vs. H1: ARCH(p) disturbance

The Ramsey RESET test and the two Breusch-Pagan tests for heteroskedasticity do not reveal any problems as we have p-values comfortably above, say, a 10% significance level. However, the White's test for homoskedasticity is rejected at a 10% significance level. Moreover, the Breusch-Godfrey LM test for autocorrelation shows some problems as the null hypothesis of no serial correlation is rejected for all the lags at the 5% significance level.

d. Model Specification

Reduce the model in successive steps by dropping insignificant variables (subjecting each step to mis-specification testing) in order to obtain a parsimonious model.

We now drop the first lag of $\ln_consump$ in the error correction model above. This gives us the following result:

Source		SS	df	MS	Number of obs =	88
Model		.004716449	5	.00094329	F(5, 82) =	13.42
Residual		.005763359	82	.000070285	Prob > F =	0.0000
Total		.010479809	87	.000120458	R-squared =	0.4501
					Adj R-squared =	0.4165
					Root MSE =	.00838

dln_consump		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
dln_consump						
L1.		-.2657825	.0869475	-3.06	0.003	-.4387488 -.0928161
L2.		.1683854	.083152	2.03	0.046	.0029696 .3338012
L3.		.2603634	.0858671	3.03	0.003	.0895463 .4311805
dln_inc		.4581345	.0802364	5.71	0.000	.2985187 .6177503
dln_inv		.0626535	.0201277	3.11	0.003	.022613 .102694
_cons		.0054381	.0028747	1.89	0.062	-.0002807 .0111569

Now we see that we do not have insignificant variables.

The mis-specification tests of this model are:

Ramsey RESET test using powers of the fitted values of dln_consump

Ho: model has no omitted variables
 F(3, 79) = 1.20
 Prob > F = 0.3151

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance
 Variables: fitted values of dln_consump

chi2(1) = 0.57
 Prob > chi2 = 0.4492

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance
 Variables: L.dln_consump L2.dln_consump L3.dln_consump dln_inc dln_inv

chi2(5) = 4.26
 Prob > chi2 = 0.5122

White's test for Ho: homoskedasticity

against Ha: unrestricted heteroskedasticity

chi2(20) = 25.95
 Prob > chi2 = 0.1675

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	25.95	20	0.1675
Skewness	9.69	5	0.0846
Kurtosis	0.62	1	0.4322
Total	36.25	26	0.0871

(n = 88)	D-H	P-value	asy.	P-value
Residuals	2.7414	0.2539	2.6979	0.2595

Breusch-Godfrey LM test for autocorrelation

lags(p)	chi2	df	Prob > chi2
1	4.309	1	0.0379
2	8.027	2	0.0181
3	9.487	3	0.0235

4		10.267	4	0.0362

H0: no serial correlation				
LM test for autoregressive conditional heteroskedasticity (ARCH)				

lags(p)		chi2	df	Prob > chi2

1		3.173	1	0.0748
2		8.125	2	0.0172
3		1.381	3	0.7100
4		0.998	4	0.9101

H0: no ARCH effects vs. H1: ARCH(p) disturbance				

We now test whether we can exclude the last lag, but the `-testparm-` command does lead to rejection, so we keep the third lag.

```
( 1)  L3.dln_consump = 0
      F( 1, 82) = 9.19
      Prob > F = 0.0032
```

e. Comparison of Models

Save the log-likelihood value of the parsimonious model (`-estimates store llr-`) and perform a likelihood ratio test comparing the model in c. with the model in d. (`-lrtest llu llr-`). You may wish to save the log-likelihood value at each step of the reduction using different names with the `estimates` command at each step.

```
. lrtest llu llr

Likelihood-ratio test                               LR chi2(1) = 0.05
(Assumption: llr nested in llu)                     Prob > chi2 = 0.8195
```

We cannot reject that the restricted and the unrestricted models are different.

f. VAR(4) Model

Estimate a VAR(4) for the three variables `dln_consump`, `dln_inc`, and `dln_inv`, that is, for the first difference of `ln_inv`, `ln_inc`, and `ln_consump`. Then determine the minimum number of lags required for the VAR, and subject it to mis-specification testing.

First we start with a VAR(4) model using the small sample correction command `-dfk-`:

```
. var dln_consump dln_inc dln_inv, lags(1/4) dfk
```

Vector autoregression

```
Sample: 1961q2 - 1982q4          No. of obs   =      87
Log likelihood = 738.3533          AIC          = -16.07709
FPE            = 2.11e-11          HQIC         = -15.63197
Det(Sigma_ml)  = 8.53e-12          SBIC         = -14.97168
```

Equation	Parms	RMSE	R-sq	chi2	P>chi2
dln_consump	13	.009863	0.3108	33.37069	0.0008
dln_inc	13	.011582	0.1728	15.4603	0.2172
dln_inv	13	.043414	0.1950	17.92655	0.1179

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
dln_consump					
dln_consump					
L1.	-.4186234	.1462402	-2.86	0.004	-.7052489 -.1319979
L2.	-.1654575	.1601944	-1.03	0.302	-.4794328 .1485178
L3.	.0699281	.148281	0.47	0.637	-.2206974 .3605536
L4.	.025482	.1316879	0.19	0.847	-.2326216 .2835856
dln_inc					
L1.	.2971221	.1203449	2.47	0.014	.0612505 .5329937
L2.	.3767172	.1288433	2.92	0.003	.1241891 .6292454
L3.	.2181382	.1315844	1.66	0.097	-.0397625 .4760389
L4.	.0940014	.1226229	0.77	0.443	-.1463351 .3343378
dln_inv					
L1.	.0043691	.0261169	0.17	0.867	-.0468191 .0555573
L2.	.0395284	.0274742	1.44	0.150	-.01432 .0933769
L3.	.0087278	.0281372	0.31	0.756	-.0464201 .0638758
L4.	-.025073	.0268227	-0.93	0.350	-.0776445 .0274986
_cons	.0076971	.0039049	1.97	0.049	.0000436 .0153507
dln_inc					
dln_consump					
L1.	.1913047	.1717152	1.11	0.265	-.145251 .5278603
L2.	-.0049856	.1881002	-0.03	0.979	-.3736552 .3636841
L3.	-.0085702	.1741116	-0.05	0.961	-.3498226 .3326822
L4.	.0246611	.154628	0.16	0.873	-.2784042 .3277263
dln_inc					
L1.	-.0722509	.1413089	-0.51	0.609	-.3492113 .2047094
L2.	.0380503	.1512877	0.25	0.801	-.2584682 .3345688
L3.	.1734287	.1545064	1.12	0.262	-.1293983 .4762557
L4.	-.0531727	.1439838	-0.37	0.712	-.3353757 .2290303
dln_inv					
L1.	.0480724	.0306665	1.57	0.117	-.0120328 .1081776
L2.	.0582109	.0322602	1.80	0.071	-.005018 .1214397
L3.	.016095	.0330387	0.49	0.626	-.0486597 .0808498

```

      L4. | -.0028715 .0314952 -0.09 0.927 -.0646011 .058858
      _cons | .011433 .0045852 2.49 0.013 .0024462 .0204198
-----+-----
dln_inv
dln_consump
      L1. | .4213112 .6436848 0.65 0.513 -.8402878 1.68291
      L2. | .4410982 .7051051 0.63 0.532 -.9408823 1.823079
      L3. | -.0088644 .6526676 -0.01 0.989 -1.288069 1.270341
      L4. | -.5482858 .5796322 -0.95 0.344 -1.684344 .5877724
      dln_inc
      L1. | .4098625 .5297049 0.77 0.439 -.6283401 1.448065
      L2. | -.1649085 .5671112 -0.29 0.771 -1.276426 .946609
      L3. | .0542709 .5791765 0.09 0.925 -1.080894 1.189436
      L4. | -.2581422 .5397318 -0.48 0.632 -1.315997 .7997127
      dln_inv
      L1. | -.2678892 .1149551 -2.33 0.020 -.4931972 -.0425813
      L2. | -.0702267 .1209292 -0.58 0.561 -.3072437 .1667902
      L3. | .1621356 .1238476 1.31 0.190 -.0806013 .4048725
      L4. | .3186896 .1180618 2.70 0.007 .0872927 .5500865
      _cons | .0071406 .0171878 0.42 0.678 -.0265469 .0408281
-----+-----
. varsoc

```

```

Selection-order criteria
Sample: 1961q2 - 1982q4      Number of obs      =      87
+-----+-----+-----+-----+-----+-----+-----+-----+
|lag |   LL   LR   df   p   FPE   AIC   HQIC   SBIC   |
+-----+-----+-----+-----+-----+-----+-----+
| 0 | 696.398                2.4e-11 -15.9402 -15.9059 -15.8552* |
| 1 | 711.682 30.568      9 0.000 2.1e-11 -16.0846 -15.9477* -15.7445 |
| 2 | 724.696 26.028      9 0.002 1.9e-11* -16.1769* -15.9372 -15.5817 |
| 3 | 729.124 8.8557      9 0.451 2.1e-11 -16.0718 -15.7294 -15.2215 |
| 4 | 738.353 18.458*     9 0.030 2.1e-11 -16.0771 -15.632 -14.9717 |
+-----+-----+-----+-----+-----+-----+-----+
Endogenous: dln_consump dln_inc dln_inv
Exogenous:  _cons

```

Althout fourth lag of ln_inv is significant, we proceed with three lags:

```
. var dln_consump dln_inc dln_inv, lags(1/3) dfk
```

Vector autoregression

```

Sample: 1961q1 - 1982q4      No. of obs      =      88
Log likelihood = 737.5558      AIC            = -16.08081
FPE           = 2.09e-11      HQIC          = -15.74057
Det(Sigma_ml) = 1.05e-11      SBIC          = -15.23627

```

```

Equation      Parms      RMSE      R-sq      chi2      P>chi2
-----
dln_consump    10      .00973    0.2953    32.68978    0.0002
dln_inc        10      .011309   0.1800    17.12733    0.0468

```

dln_inv 10 .045251 0.1096 9.599902 0.3838

		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
dln_consump							
dln_consump							
	L1.	-.3930788	.1385722	-2.84	0.005	-.6646753	-.1214823
	L2.	-.1281016	.1441055	-0.89	0.374	-.4105431	.15434
	L3.	.1015717	.1278624	0.79	0.427	-.149034	.3521773
dln_inc							
	L1.	.2940007	.1174905	2.50	0.012	.0637236	.5242778
	L2.	.3438497	.1205479	2.85	0.004	.1075802	.5801192
	L3.	.1810283	.1187276	1.52	0.127	-.0516734	.4137301
dln_inv							
	L1.	.0034602	.0256856	0.13	0.893	-.0468826	.0538031
	L2.	.0437956	.0264529	1.66	0.098	-.0080512	.0956423
	L3.	.0168297	.0263789	0.64	0.523	-.034872	.0685314
_cons		.0091472	.0034369	2.66	0.008	.0024111	.0158833
dln_inc							
dln_consump							
	L1.	.1923453	.161054	1.19	0.232	-.1233148	.5080055
	L2.	-.0339177	.167485	-0.20	0.840	-.3621824	.294347
	L3.	-.0434862	.1486067	-0.29	0.770	-.33475	.2477776
dln_inc							
	L1.	-.0815851	.136552	-0.60	0.550	-.3492222	.186052
	L2.	.0624807	.1401055	0.45	0.656	-.212121	.3370824
	L3.	.1984851	.1379898	1.44	0.150	-.07197	.4689402
dln_inv							
	L1.	.0476297	.0298528	1.60	0.111	-.0108808	.1061401
	L2.	.0598946	.0307446	1.95	0.051	-.0003637	.1201529
	L3.	.0155151	.0306586	0.51	0.613	-.0445747	.0756048
_cons		.011234	.0039945	2.81	0.005	.003405	.019063
dln_inv							
dln_consump							
	L1.	.4585587	.644436	0.71	0.477	-.8045127	1.72163
	L2.	.4443132	.6701688	0.66	0.507	-.8691936	1.75782
	L3.	-.199181	.5946297	-0.33	0.738	-1.364634	.9662718
dln_inc							
	L1.	.3711541	.5463946	0.68	0.497	-.6997597	1.442068
	L2.	.2534399	.5606132	0.45	0.651	-.8453418	1.352222
	L3.	.3569366	.5521478	0.65	0.518	-.7252532	1.439126
dln_inv							
	L1.	-.2598316	.1194521	-2.18	0.030	-.4939533	-.0257098
	L2.	-.1253296	.1230204	-1.02	0.308	-.3664451	.115786
	L3.	.0466697	.1226763	0.38	0.704	-.1937714	.2871107

```

      |
      _cons | -.0100135   .0159833   -0.63   0.531   -.0413401   .0213131
-----+-----
. varsoc

      Selection-order criteria
      Sample: 1961q1 - 1982q4
      Number of obs   =      88
+-----+-----+
|lag |   LL   LR   df   p   FPE   AIC   HQIC   SBIC |
+-----+-----+
| 0 | 703.801                2.4e-11 -15.9273 -15.8933 -15.8428* |
| 1 | 719.587 31.572   9 0.000 2.1e-11 -16.0815 -15.9454* -15.7437 |
| 2 | 732.981 26.787*  9 0.002 1.9e-11* -16.1814* -15.9432 -15.5902 |
| 3 | 737.556 9.1506  9 0.423 2.1e-11 -16.0808 -15.7406 -15.2363 |
+-----+-----+
Endogenous:  dln_consump dln_inc dln_inv
Exogenous:   _cons

```

The -varsoc- command shows us that we can drop the third lag, so we try with two lags:

```

. dln_consump dln_inc dln_inv, lags(1/2) dfk

Vector autoregression

Sample: 1960q4 - 1982q4
Log likelihood = 742.2131
FPE = 1.84e-11
Det(Sigma_ml) = 1.15e-11
No. of obs = 89
AIC = -16.20704
HQIC = -15.97035
SBIC = -15.61983

Equation      Parms      RMSE      R-sq      chi2      P>chi2
-----+-----+-----+-----+-----+-----+
dln_consump      7      .009938      0.2400      25.88962      0.0002
dln_inc          7      .011224      0.1514      14.62996      0.0233
dln_inv          7      .044295      0.1051      9.633771      0.1409
-----+-----+-----+-----+-----+-----+

-----+-----+-----+-----+-----+-----+
|      Coef.   Std. Err.      z    P>|z|      [95% Conf. Interval]
-----+-----+-----+-----+-----+-----+
dln_consump |
dln_consump |
L1. | -.2845172   .1274068   -2.23   0.026   -.5342299   -.0348044
L2. | -.1159776   .1270233   -0.91   0.361   -.3649386   .1329834
|
dln_inc |
L1. | .2893204   .112313   2.58   0.010   .069191   .5094497
L2. | .3664341   .1089872   3.36   0.001   .1528231   .5800451
|
dln_inv |
L1. | .0027381   .0255556   0.11   0.915   -.0473499   .0528261
L2. | .0497402   .025462   1.95   0.051   -.0001644   .0996447
|
_cons | .0123795   .0029602   4.18   0.000   .0065776   .0181813
-----+-----+-----+-----+-----+-----+
dln_inc |

```



```

dln_consump |
    L1. | .3050571 .1438994 2.12 0.034 .0230195 .5870947
    L2. | .0490208 .1434662 0.34 0.733 -.2321677 .3302094
    |
dln_inc |
    L1. | -.1232543 .1268517 -0.97 0.331 -.371879 .1253704
    L2. | .0209769 .1230954 0.17 0.865 -.2202857 .2622394
    |
dln_inv |
    L1. | .0433473 .0288637 1.50 0.133 -.0132244 .0999191
    L2. | .0616319 .028758 2.14 0.032 .0052673 .1179965
    |
_cons | .0125949 .0033434 3.77 0.000 .006042 .0191478
-----+-----
dln_inv
dln_consump |
    L1. | .6520473 .5678885 1.15 0.251 -.4609936 1.765088
    L2. | .5980687 .566179 1.06 0.291 -.5116217 1.707759
    |
dln_inc |
    L1. | .3374819 .5006109 0.67 0.500 -.6436975 1.318661
    L2. | .1827302 .4857871 0.38 0.707 -.7693951 1.134855
    |
dln_inv |
    L1. | -.2725654 .1139085 -2.39 0.017 -.4958218 -.0493089
    L2. | -.1340503 .1134912 -1.18 0.238 -.3564891 .0883884
    |
_cons | -.0099191 .0131944 -0.75 0.452 -.0357798 .0159415
-----+-----
. varsoc

Selection-order criteria
Sample: 1960q4 - 1982q4 Number of obs = 89
+-----+-----+
|lag | LL LR df p FPE AIC HQIC SBIC |
+-----+-----+
| 0 | 712.092 2.4e-11 -15.9346 -15.9008 -15.8508* |
| 1 | 728.167 32.15 9 0.000 2.1e-11 -16.0936 -15.9584 -15.7581 |
| 2 | 742.213 28.092* 9 0.001 1.8e-11* -16.207* -15.9704* -15.6198 |
+-----+-----+
Endogenous: dln_consump dln_inc dln_inv
Exogenous: _cons
. varlmar

Lagrange-multiplier test
+-----+-----+
| lag | chi2 df Prob > chi2 |
+-----+-----+
| 1 | 8.8693 9 0.44942 |
| 2 | 10.9722 9 0.27762 |
+-----+-----+
H0: no autocorrelation at lag order
. varlmar

Jarque-Bera test
+-----+-----+
| Equation | chi2 df Prob > chi2 |

```

dln_consump	10.245	2	0.00596	
dln_inc	1.145	2	0.56398	
dln_inv	7.029	2	0.02976	
ALL	18.419	6	0.00527	

Skewness test

Equation	Skewness	chi2	df	Prob > chi2
dln_consump	-.73238	7.956	1	0.00479
dln_inc	-.10013	0.149	1	0.69975
dln_inv	.32757	1.592	1	0.20709
ALL	9.697	3	0.02133	

Kurtosis test

Equation	Kurtosis	chi2	df	Prob > chi2
dln_consump	3.7855	2.288	1	0.13036
dln_inc	2.4816	0.997	1	0.31810
dln_inv	4.2109	5.437	1	0.01971
ALL	8.722	3	0.03322	

dfk estimator used in computations

. varwle

Equation: dln_consump

lag	chi2	df	Prob > chi2
1	7.675739	3	0.053
2	17.32727	3	0.001

Equation: dln_inc

lag	chi2	df	Prob > chi2
1	9.628048	3	0.022
2	5.914459	3	0.116

Equation: dln_inv

lag	chi2	df	Prob > chi2
1	7.873887	3	0.049
2	2.993971	3	0.393

Equation: All

lag	chi2	df	Prob > chi2
-----	------	----	-------------

```

| 1 | 44.06615 9 0.000 |
| 2 | 28.28163 9 0.001 |
+-----+

```

g. Final Model

Finally, drop individual variables from each of the three equations until you end up with a parsimonious model.

We now drop individual variables from each equation until we end up with the following parsimonious model with eight constraints:

```

. var dln_consump dln_inc dln_inv, lags(1/2) dfk constraint (1 2 3 4 5 6 7 8)
Estimating VAR coefficients

```

```

Iteration 1: tolerance = .08076175
Iteration 2: tolerance = .00197295
Iteration 3: tolerance = .0000493
Iteration 4: tolerance = 1.529e-06
Iteration 5: tolerance = 5.099e-08

```

Vector autoregression

```

Sample: 1960q4 - 1982q4          No. of obs   =      89
Log likelihood = 739.5869        AIC          = -16.14802
FPE           = 2.18e-15        HQIC       = -15.91133
Det(Sigma_ml) = 1.63e-15        SBIC       = -15.56081

```

Equation	Parms	RMSE	R-sq	chi2	P>chi2
dln_consump	5	.009606	0.2292	37.83738	0.0000
dln_inc	4	.01085	0.1394	15.67358	0.0013
dln_inv	4	.042994	0.0849	9.673415	0.0216

```

( 1) [dln_consump]L.dln_inv = 0
( 2) [dln_inc]L2.dln_inc = 0
( 3) [dln_inv]L2.dln_inc = 0
( 4) [dln_inc]L2.dln_consump = 0
( 5) [dln_inv]L.dln_inc = 0
( 6) [dln_inc]L.dln_inc = 0
( 7) [dln_inv]L2.dln_inv = 0
( 8) [dln_consump]L2.dln_consump = 0

```

		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
dln_consump							
dln_consump							
	L1.	-.2560975	.1075375	-2.38	0.017	-.4668671	-.0453279
	L2.	2.70e-18	1.79e-17	0.15	0.880	-3.24e-17	3.78e-17
dln_inc							
	L1.	.2899923	.0825004	3.52	0.000	.1282945	.4516901


```

+-----+
|          dln_consump |          11.036  2   0.00401 |
|          dln_inc     |          0.869  2   0.64764 |
|          dln_inv     |         12.553  2   0.00188 |
|          ALL         |         24.457  6   0.00043 |
+-----+

Skewness test
+-----+
|          Equation | Skewness  chi2  df  Prob > chi2 |
+-----+
|          dln_consump | -.74579   8.250  1   0.00407 |
|          dln_inc     | -.17547   0.457  1   0.49915 |
|          dln_inv     | .42156    2.636  1   0.10446 |
|          ALL         | 11.343    3     0.01001 |
+-----+

Kurtosis test
+-----+
|          Equation | Kurtosis  chi2  df  Prob > chi2 |
+-----+
|          dln_consump | 3.8666    2.785  1   0.09514 |
|          dln_inc     | 2.6666    0.412  1   0.52091 |
|          dln_inv     | 4.6353    9.917  1   0.00164 |
|          ALL         | 13.114    3     0.00440 |
+-----+

dfk estimator used in computations
. varstable, graph

Eigenvalue stability condition
+-----+
|          Eigenvalue |          Modulus |
+-----+
|          .6289397   |          .62894  |
|          -.4992968   |          .499297  |
|          -.3415027 + .3613781i |          .49721  |
|          -.3415027 - .3613781i |          .49721  |
|          .02051253 + .4183026i |          .418805  |
|          .02051253 - .4183026i |          .418805  |
+-----+

All the eigenvalues lie inside the unit circle.
VAR satisfies stability condition.

```

The likelihood-ratio test above shows us however that we cannot reject that the unrestricted model from section f. on page 20 is different from the restricted model just estimated. The mis-specification tests however portrays some problems as we for instance cannot reject that we have no autocorrelation. Finally, the last part shows the stability analysis, which concludes that all the eigenvalues lie inside the unit circle:

